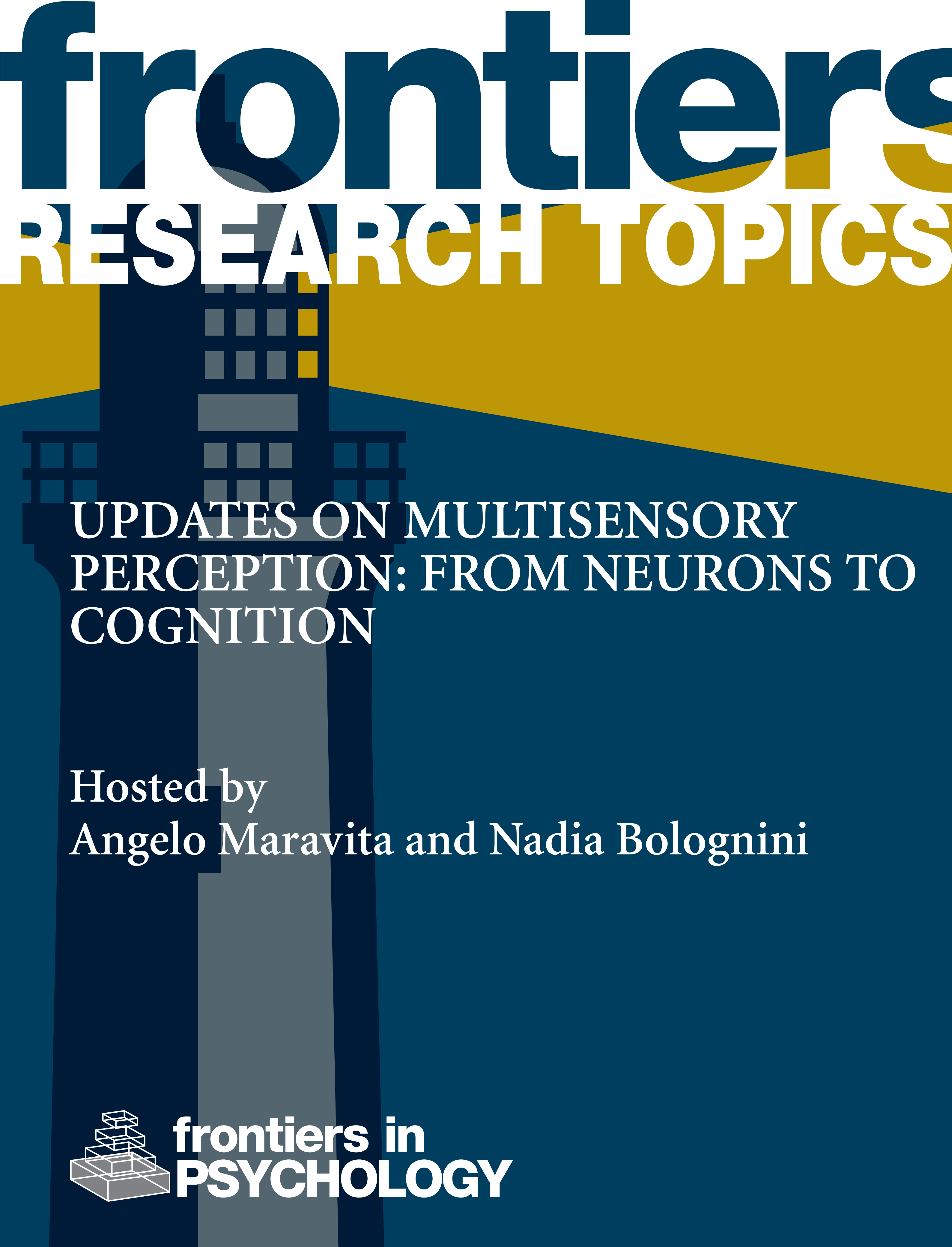

frontiers
RESEARCH TOPICS

# UPDATES ON MULTISENSORY PERCEPTION: FROM NEURONS TO COGNITION

Hosted by
Angelo Maravita and Nadia Bolognini

frontiers in
PSYCHOLOGY

### FRONTIERS COPYRIGHT STATEMENT

Cover image provided by Ibbl sarl, Lausanne CH

**ISSN** 1664-8714
**ISBN** 978-2-88919-058-4
**DOI** 10.3389/978-2-88919-058-4

## ABOUT FRONTIERS

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## FRONTIERS JOURNAL SERIES

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing.

All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## DEDICATION TO QUALITY

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## WHAT ARE FRONTIERS RESEARCH TOPICS?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area!

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

# UPDATES ON MULTISENSORY PERCEPTION: FROM NEURONS TO COGNITION

Hosted By
**Angelo Maravita**, University of Milano–Bicocca, Italy
**Nadia Bolognini**, University of Milano–Bicocca, Italy

In recent years there has been a dramatic progress in understanding how stimuli from different sensory modalities are integrated among each other. Multisensory integration results in a unitary representation of the world that strongly characterizes perception and cognition in humans.

Knowledge about multi sensory integration has research techniques and approaches, including neurophysiology, experimental psychology, neuropsychology, neuroimaging, and computational modelling. This special issue aims at presenting an up-to-date integrative overview of the physiological, psychological, developmental, and functional processes associated with multisensory integration.

The proposed collection of papers is organized thematically into sections, each featuring a state-of-the-art review of key themes in multisensory research, from more approaches in the animal, to the study of multisensory perception and cognition in humans. Specifically, this special issue will consider: The physiological mechanisms of multisensory processing in cortical and subcortical brain structures of model animal species, (rat, cat, and monkey); current biologically inspired computational modelling of multisensory integration; evidence about the multisensory contributions to perception in humans, as highlighted by psychophysical and neuropsychological evidence; the neural basis of multisensory processing in the human brain uncovered by recent neuroimaging techniques, including EEG, PET, fMRI; the consequences of the breakdown of normal sensory integration as shown by studies with techniques of brain stimulation in humans; developmental processes of multisensory perception in humans and the constrains for the emergence of multisensory processes in relation to sensory experience; the issue of crossmodal neuroplasticity concerning behavioral and neural changes following sensory deprivation.

The challenge of this Research Topic is to provide an interdisciplinary context allowing to understand the basic principles of multisensory integration in humans and the key issues that this fascinating field of study rises for future research.

# Table of Contents

frontiers in
PSYCHOLOGY

EDITORIAL
published: 03 May 2012
doi: 10.3389/fpsyg.2012.00122

# Interactions between senses: updating on neural mechanisms and behavioral evidence

**Nadia Bolognini* and Angelo Maravita**

*Department of Psychology, University of Milano Bicocca, Milano, Italy*
**Correspondence: nadia.bolognini@unimib.it*

In recent years there has been a dramatic progress in understanding how stimuli from different sensory modalities are integrated among each other. Multisensory integration results in a unitary representation of the world that strongly characterizes perception and cognition in humans. The body of knowledge acquired so far on multisensory integration has been gained through several research techniques and approaches, including neurophysiology, experimental psychology, neuropsychology, neuroimaging, and computational modeling. This special issue aims at presenting an up-to-date overview of the research on multisensory integration. In particular, the proposed collection of papers features state-of-the-art reviews or original articles about key themes in multisensory research, considering novel evidence on the physiological mechanism of multisensory integration at cell level, and on the behavioral effects of multisensory integration on perception and cognition in humans.

The issue starts with novel findings about the functional properties of multisensory integration in multisensory neurons of the superior colliculus (SC). Perrault et al. (2011) show how the state of the multisensory cells in SC and their integrative principles are modulated as a consequence of the exposure to a series of testing stimuli during an experiment, following the well-known "inverse effectiveness principle." The evidence provided has both empirical and practical implications: first, it shows that multisensory integration in the SC is highly plastic and multisensory neurons can adapt to rapidly changing environmental events; second, the neuron studied at the beginning of an experiment is not the same at the end of it, and this should represent an important caveat in the interpretation of experimental data.

The importance of biologically inspired computational modeling of multisensory integration is well highlighted by Cuppini et al. (2011) who describe two neural network models. The first one is aimed at reproducing and explaining the main physiological features of multisensory integration in SC's neurons, and its development during postnatal life depending on sensory experience. The second model tackles the problem of how tactile stimuli on a body part and visual (or auditory) stimuli close to the same body part are integrated in multimodal parietal neurons to form the perception of near peripersonal space. The proposed models are helpful for the interpretation and the integration of behavioral and physiological evidence, and they offer a novel perspective for guiding future experiments.

The issue of the effects of visual–auditory interactions on human perception is the topic of four behavioral experiments, and a review article. Romei et al. (2011) nicely demonstrate that visual duration discrimination can be enhanced by temporally congruent auditory stimuli, through changes in visual sensitivity. Alink et al. (2012) show that moving sounds can capture the direction of visual stimuli, especially when visual motion direction is ambiguous. Arrighi et al. (2011) focus on the effect of the attentional resources on crossmodal perceptual abilities, and show that vision and audition rely on separate attentional resources. Finally, Chen et al. (2011) provide novel evidence for the effect of the auditory semantic context on visual awareness by using the binocular rivalry paradigm. All together, these research findings provide original support to the view that multisensory (audio–visual) experience can impact unisensory processing. This proposal is deeply discussed by Shams et al. (2011). The authors review three different types of learning, namely perceptual learning, sensory recalibration, and associative learning. In all these cases, exposure to multisensory stimuli profoundly influences the subsequent unisensory processing, suggesting that the modification of unisensory representations by multisensory relationships may be a general learning strategy employed by the brain.

Visual–tactile interactions are investigated by two studies, and discussed in a review article. First, Gori et al. (2011) explore the mechanisms mediating crossmodal facilitation and summation between visual and tactile motion. The authors find that while summation produced a generic, non-specific improvement of thresholds, probably reflecting higher-order interaction of decision signals, facilitation reveals a strong, direction-specific interaction, likely reflecting low-level sensory interactions. Mancini et al. (2011) adopt a neuropsychological approach to investigate the integration between vision and haptics by adopting an original bisection task of a crossmodal illusion, the Judd variant of the Müller-Lyer illusion. The main finding is that, although a right-hemisphere damage my induce modality-specific deficits of spatial attention and representation, the ability of properly integrate visual and tactile sensory inputs is spared in right-brain-damaged patients. This evidence is in line with another type of crossmodal phenomenon that is maintained in right-brain-damaged patients, namely crossmodal extinction. Jacobs et al. (2011) provide a critical review of the research on crossmodal extinction, demonstrating how the study of this phenomenon has critically contributed to increase our understanding of how the integration of stimuli perceived in multiple sensory modalities is used by the human brain to build coherent representations of the space that directly surrounds us.

Moving to complex cognitive implications of sensory processing, Forgiarini et al. (2011) show the existence of a racial bias in the autonomic reaction to other people's pain and its link with implicit racial biases as assessed through the well-known implicit association test. The authors show and discuss the role of others' race in moderating empathic reactions, which represents a crucial clue for understanding to what extent social interactions may be influenced by deeply rooted automatic and uncontrollable processes.

The multisensory and sensorimotor basis of bodily self-consciousness is the topic of the review by Ionta et al. (2011), which focus on the multisensory role of the temporo-parietal junction in bodily self-consciousness, as highlighted by the study of neurological patients suffering from out-of-body experiences, and of healthy subjects undergoing experimentally induced multisensory conflicts.

This special issue ends with a review by Bolognini and Maravita (2011) that critically discuss why techniques of non-invasive brain stimulation can represent a unique and powerful approach to inform models of causal relations between specific brain regions and multisensory functions.

Overall, the challenge of this special issue is to provide an interdisciplinary context allowing to understand the basic principles of multisensory integration in animals and humans and the key issues that this fascinating field of study rises for future research. To conclude, we would like to deeply thank the many scientists who contributed so nicely to this special issue.

## REFERENCES

Alink, A., Euler, F., Galeano, E., Krugliak, A., Singer, W., and Kohler, A. (2012). Auditory motion capturing ambiguous visual motion. *Front. Psychol.* 2:391. doi: 10.3389/fpsyg.2011.00391

Arrighi, R., Lunardi, R., and Burr, D. (2011). Vision and audition do not share attentional resources in sustained tasks. *Front. Psychol.* 2:56. doi: 10.3389/fpsyg.2011.00056

Bolognini, N., and Maravita, A. (2011). Uncovering multisensory processing through non-invasive brain stimulation. *Front. Psychol.* 2:46. doi: 10.3389/fpsyg.2011.00046

Chen, Y., Yeh, S., and Spence, C. (2011). Crossmodal constraints on human perceptual awareness: auditory semantic modulation of binocular rivalry. *Front. Psychol.* 2:212. 10.3389/fpsyg.2011.00212

Cuppini, C., Magosso, E., and Ursino, M. (2011). Organization, maturation, and plasticity of multisensory integration: insights from computational modeling studies. *Front. Psychol.* 2:77. doi: 10.3389/fpsyg.2011.00077

Forgiarini, M., and Gallucci, M., and Maravita, A. (2011). Racism and the empathy for pain on our skin. *Front. Psychol.* 2:108. doi: 10.3389/fpsyg.2011.00108

Gori, M., Mazzilli, G., Sandini, G., and Burr, D. (2011). Cross-sensory facilitation reveals neural interactions between visual and tactile motion in humans. *Front. Psychol.* 2:55. doi: 10.3389/fpsyg.2011.00055

Ionta, S., Gassert, R., and Blanke, O. (2011). Multi-sensory and sensorimotor foundation of bodily self-consciousness–an interdisciplinary approach. *Front. Psychol.* 2:383. doi: 10.3389/fpsyg.2011.00383

Jacobs, S., Brozzoli, C., Hadj-Bouziane, F., Meunier, M., and Farnè, A. (2011). Studying multisensory processing and its role in the representation of space through pathological and physiological crossmodal extinction. *Front. Psychol.* 2:89. doi: 10.3389/fpsyg.2011.00089

Mancini, F., Bricolo, E., Mattioli, F. C., and Vallar, G. (2011). Visuo-haptic interactions in unilateral spatial neglect: the cross modal Judd illusion. *Front. Psychol.* 2:341. doi: 10.3389/fpsyg.2011.00341

Perrault, T. J. Jr., Stein, B. E., and Rowland, B. A. (2011). Non-stationarity in multisensory neurons in the superior colliculus. *Front. Psychol.* 2:144. doi: 10.3389/fpsyg.2011.00144

Romei, V., DeHaas, B., Mok, R. M., and Driver, J. (2011). Auditory stimulus timing influences perceived duration of co-occurring visual stimuli. *Front. Psychol.* 2:215. doi: 10.3389/fpsyg.2011.00215

Shams, L., Wozny, D. R., Kim, R., and Seitz, A. (2011). Influences of multisensory experience on subsequent unisensory processing. *Front. Psychol.* 2:264. doi: 10.3389/fpsyg.2011.00264

*Received: 22 March 2012; accepted: 07 April 2012; published online: 03 May 2012.*

*Citation: Bolognini N and Maravita A (2012) Interactions between senses: updating on neural mechanisms and behavioral evidence. Front. Psychology **3**:122. doi: 10.3389/fpsyg.2012.00122*

*This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.*

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 04 July 2011
doi: 10.3389/fpsyg.2011.00144

# Non-stationarity in multisensory neurons in the superior colliculus

**Thomas J. Perrault Jr.*, Barry E. Stein and Benjamin A. Rowland**

*Department of Neurobiology and Anatomy, Wake Forest University School of Medicine, Winston Salem, NC, USA*

***Edited by:***
*Nadia Bolognini, University of Milano-Bicocca, Italy*

***Reviewed by:***
*Angelo Maravita, University of Milano-Bicocca, Italy*
*Cristiano Cuppini, University of Bologna, Italy*

****Correspondence:***
*Thomas J. Perrault, Department of Neurobiology and Anatomy, Wake Forest University School of Medicine, Medical Center Boulevard, Winston Salem, NC 27157, USA.*
*e-mail: tperraul@wfubmc.edu*

The superior colliculus (SC) integrates information from multiple sensory modalities to facilitate the detection and localization of salient events. The efficacy of "multisensory integration" is traditionally measured by comparing the magnitude of the response elicited by a cross-modal stimulus to the responses elicited by its modality-specific component stimuli, and because there is an element of randomness in the system, these calculations are made using response values averaged over multiple stimulus presentations in an experiment. Recent evidence suggests that multisensory integration in the SC is highly plastic and these neurons adapt to specific anomalous stimulus configurations. This raises the question whether such adaptation occurs during an experiment with traditional stimulus configurations; that is, whether the state of the neuron and its integrative principles are the same at the beginning and end of the experiment, or whether they are altered as a consequence of exposure to the testing stimuli even when they are pseudo-randomly interleaved. We find that unisensory and multisensory responses do change during an experiment, and that these changes are predictable. Responses that are initially weak tend to potentiate, responses that are initially strong tend to habituate, and the efficacy of multisensory integration waxes or wanes accordingly during the experiment as predicted by the "principle of inverse effectiveness." These changes are presumed to reflect two competing mechanisms in the SC: potentiation reflects increases in the expectation that a stimulus will occur at a given location relative to others, and habituation reflects decreases in stimulus novelty. These findings indicate plasticity in multisensory integration that allows animals to adapt to rapidly changing environmental events while suggesting important caveats in the interpretation of experimental data: the neuron studied at the beginning of an experiment is not the same at the end of it.

**Keywords: multisensory, superior colliculus**

## INTRODUCTION

There is a significant difference in the way that we think about the brain and the way in which we conduct experiments and analyze data. While we appreciate that the brain is malleable and adapts to experience, we know that neural activity can be random, and we draw our conclusions from analyses that require data to be averaged over multiple trials. In the study of sensory neurophysiology, we commonly attempt to attenuate adaptation to different stimuli by randomly interleaving their presentation. But is the neuron at the beginning of the experiment in the same state at the end given these efforts?

We found that, in the case of superior colliculus (SC) neurons, it is not. The SC is unique in that it is capable of integrating information from multiple sensory modalities (Stein and Arigbede, 1972; Stein and Dixon, 1979; Meredith and Stein, 1983, 1986; King and Palmer, 1985; Meredith et al., 1992; Peck et al., 1993; Stein and Meredith, 1993; Wallace et al., 1993, 1996; Wallace and Stein, 2001; Populin and Yin, 2002; Perrault et al., 2003; Alvarado et al., 2008, 2009; Zahar et al., 2009). The SC contains multisensory neurons that receive unisensory signals from independent channels and integrate that information in the form of response enhancements and depressions. This integrative capacity makes them well suited for stimulus detection and localization (Jay and Sparks, 1987; Lee et al., 1988; Frens and Van Opstal, 1998; Jiang et al., 2002; Burnett et al., 2004; Bell et al., 2005). The most common method to assess these enhancements and depressions employ techniques that evaluate the averaged activity of these neurons to both modality-specific and cross-modal stimuli over many trials. While integrative responses are clearly demonstrated throughout the literature, it is also apparent that neurons within the SC exhibit substantial variation in activity levels across the population as well as within the responsivity of a single neuron. We sought to explore how the changes in an individual neuron's responsivity might be explained. Our fundamental hypothesis is that it reflects two opposing forces, one that promotes the detection of weakly effective stimuli, and another that demotes responses to strongly effective stimuli. We predicted that there would be fluctuations in the responses of SC neurons during an experiment, with the specific prediction that strong responses would decrease in magnitude ("habituate"), while weak responses would increase ("potentiate").

In an analysis of an extensive dataset ($n = 110$) in which multisensory neurons in the SC were presented with visual, auditory, and visual–auditory combinations of stimuli, we found significant changes in their response magnitudes that were consistent with our principal hypothesis (i.e., strong responses habituated, weak responses potentiated). These trends were not due to random fluctuations. As a consequence, the "benefit" provided by integrating signals across sensory channels ("multisensory integration") changed during the course of the experiment in a manner consistent with the "principle of inverse effectiveness" (Meredith et al., 1987; Stein and Stanford, 2008). The potency of multisensory integration is therefore not constant in time either during a response (Rowland et al., 2007), or over the course of an experiment.

Instead, unisensory and multisensory responses are not stationary, an observation that is consistent with recent findings (see Yu et al., 2009). However, unlike that study, the dataset analyzed here was recorded under experimental conditions designed to attenuate such changes. Predictable changes occurred despite these efforts, thereby revealing another aspect by which inherent neural plasticity can have a substantial impact on the way in which we interpret data.

## MATERIALS AND METHODS

All procedures were carried out in accordance with the National Institutes of Health guidelines for animal research and were in compliance with an approved protocol at the Wake Forest University School of Medicine, which is accredited by the American Association for the Accreditation of Laboratory Animal Care. Experiments were performed in three adult cats weighing 2.5–5.0 kg. Animals were prescreened for normal vision and hearing prior to inclusion in the study.

### IMPLANTATION PROCEDURE

An initial surgical procedure was performed placing a recording well/head-holding device on the skull prior to any electrophysiology recordings (McHaffie and Stein, 1983). Animals were initially rendered tractable with an intramuscular injection of ketamine HCl (20 mg/kg) and acepromazine maleate (0.2–0.4 mg/kg). A steady plane of anesthesia maintained using isoflurane (1–4%) following endotracheal intubation. Throughout the surgery, hydration was maintained with intravenous infusion of lactated Ringer solution (4–8 ml/h) via the saphenous vein. This was followed by postsurgical subcutaneous administration of lactated Ringer (30 ml/kg). Expiratory $CO_2$ and temperature were monitored to remain within normal limits. Eyes were protected with constant application of sterile artificial tears to prevent corneal drying during the duration of the procedure. Once anesthetized, the animal was placed in a stereotaxic frame, and a craniotomy exposed the cortex overlying the SC. A stainless steel chamber that provides access to the SC as well as holds the head was affixed to the skull using stainless steel bone screws and orthopedic bone cement. Analgesics (butorphanol 0.1–0.4 mg/kg or ketoprofen 1–2 mg/kg) were given as needed during recovery.

### GENERAL RECORDING PROCEDURE

During electrophysiological recordings the animal was rendered tractable with an intramuscular injection of ketamine (20 mg/kg) and acepromazine (0.4 mg/kg). Animals were intubated and stabilized in a head holder without wounds or pressure points. A cannula was placed in the saphenous vein for the continuous delivery of anesthetic (ketamine: 4–8 mg/kg/h), paralytic (pancuronium bromide: 0.2 mg/kg/h), and fluids (lactated Ringer: 4–8 ml/h). Paralysis and artificial respiration are necessary because eye movements can produce significant displacement of visual receptive fields. Maintenance of adequate levels of anesthesia was done by monitoring multiple vital signs, including expiratory $CO_2$, heart rate, and blood pressure. For purposes of receptive field mapping, the pupils were dilated with 1% atropine sulfate and corrective contact lenses were placed on the anesthetized (0.5% proparacaine hydrochloride ophthalmic solution) corneas to adjust for retinoscopically determined refractive errors. The optic disks were rear-projected and focused onto a 91-cm-diameter translucent hemisphere placed 45 cm from the eyes. Receptive field maps acquired from multiple sessions in the same and different animals could then be registered by aligning the position of the optic disk. Individual animals underwent recording sessions 1–3 times per week. Anesthesia and paralytic was reversed and on return of normal respiration and locomotion, the animal was returned to its home cage. Experiments generally lasted between 8 and 12 h.

### NEURONAL ISOLATION AND RECORDING

Tungsten microelectrodes (tip diameter: 1–3 μm, impedance: 1–3 MΩ at 1 kHz) were positioned with a Kopf micromanipulator and lowered into the intermediate layers of the SC. Position was confirmed by the characteristic visual activity elicited by the superficial layers. The electrode was advanced into stratum opticum (the transitional layer between the superficial and deep SC) by means of a hydraulic microdrive. From here, the electrode was advanced in 10-μm steps while presenting visual and auditory search stimuli as in previous studies (Meredith and Stein, 1986; Wallace et al., 1993). Single units were isolated (criterion signal: noise = 3:1) and digitized by means of a window discriminator (FHC). Neural activity was amplified and monitored, and data were collected using a customized suite of software that employs the 1401 Plus data acquisition system (Cambridge Electronic Design). In the current study, only visual–auditory multisensory neurons were examined.

### SENSORY STIMULI

Stationary visual stimuli consisted of the 50- to 100-ms illumination of a light-emitting diode (LED; 660 nm λ) placed within the receptive field (see Receptive Field Mapping). Moving visual stimuli consisted of slits, bars, or spots of light projected onto the translucent hemisphere, the movement speed, amplitude, and direction of which could be independently controlled. Whereas the intensity of stationary stimuli was computer controlled, the intensity of moving stimuli was controlled using neutral density filters. In both circumstances stimulus intensity ranged from 0.11 to 13.0 cd/m$^2$ with a background luminance of 0.10 cd/m$^2$. Auditory stimuli were delivered in a free-field setting and consisted of 50- to 100-ms duration broadband noise bursts (20 Hz–10 kHz). These stimuli were digitally synthesized and delivered through speakers that could be positioned at any location in auditory space. Auditory stimulus intensities ranged from 50.6 to 70.0 dB sound pressure level (SPL) against a background SPL of 50.0 dB. Visual and auditory intensities used for testing were determined by presenting a range of intensities that elicited a threshold response as well as a saturated response. Once this was determined visual and auditory stimuli were matched to create the cross-modal pair based upon their relative position along their unisensory saturation curve. Modality-specific and cross-modal stimuli were then presented randomly for a minimum of 15 trials at the most sensitive location within their respective excitatory receptive field. For all trials, neuronal activity was recorded for 2–3 s with a 500 ms interval prior to stimulus presentation. An interstimulus interval ranging from 5 to 10 s was used for each trial.

### RECEPTIVE FIELD MAPPING

The borders of each visual receptive field were mapped onto the translucent hemisphere by moving the optimum stimulus, projected from a handheld pantoscope, from the periphery inward from all directions until an enclosed responsive area was defined. Auditory receptive fields were mapped using brief (50 ms) broadband noise bursts delivered from a speaker that could be positioned at any location on a hoop that could be freely rotated about the animal's interaural axis. The typical steps in speaker location represented ~15° of auditory angle in both the azimuthal and elevation dimensions. The location of the stimulus was randomly varied, and a positive response (i.e., a location within the receptive field) was one in which the stimulus-evoked response was readily discernible above background activity. For purposes of receptive field mapping, auditory stimuli were 15 dB above the neuron's previously determined threshold. Receptive fields were transposed from the hemisphere and plotted on standardized representations of visual and auditory space.

### DATA ANALYSIS

Response magnitude was identified by first demarcating the onset and offset of the response using a three-step geometric method as in the past (Rowland et al., 2007). This "response window" was calculated using data from all trials from a particular stimulus presentation condition. The window of time 500 ms prior to stimulus onset was used to calculate the spontaneous rate. We then counted the number of impulses in the response window on each trial and subtracted the expected number given the size of the window and the spontaneous rate, producing a trial-by-trial estimate of response magnitude. A simple linear regression was used to determine the slope of the response magnitude vs. trial number trend.

Because it was possible, in principle, for any changes to be due to random fluctuations, the analysis had multiple levels. One would expect, based on random fluctuations, that responses that were randomly large on the first trial would be smaller on the last, and responses that were randomly small on the first trial would be larger on the last ("regression to the mean"). To compensate, in our analysis we examined how predictive the averages of the response magnitudes on the first few trials were of the overall trend. If averaging several initial trials still provided a better estimate relative to averaging all trials, then regression to the mean would be a less likely inference. This was tested with correlation values.

We then examined how predictive the response to a stimulus would be of the response when it was presented a second, third, or fourth time in the randomly interleaved series (i.e., with intervening stimuli). In a circumstance in which neural responses are random, the correlation coefficient should be randomly distributed in each of these comparisons. However, if the neural response is changing in a predictable way, then the first response should be a better predictor of the response to the second presentation than the third presentation. Again, correlation was used to quantify these relationships.

Finally, we compared the slopes of the changes in the multisensory responses to the best (i.e., largest) unisensory responses. If multisensory responses reflect fluctuations in the unisensory input magnitudes, there should be a correlation between these values according to the principle of inverse effectiveness. On the other hand, if the results were due to randomness, there would be no correlation: unisensory efficacy and multisensory efficacy would be unrelated in this comparison.

## RESULTS

The primary observation was that initially strong responses tended to get weaker, while weak responses got stronger, during the course of the experiment. This occurred for both unisensory and multisensory responses even when stimulus presentations were randomly interleaved in an attempt to suppress such changes. The results did not reflect random changes in the neuron's responsiveness, because changes could be better predicted by averaging the responses on a few initial trials, response magnitudes on a given trial were better predicted by more proximate responses, and there was a good correlation between the changes observed in the multisensory and unisensory responses. As a consequence, the efficacy of multisensory integration during the course of an experiment changed in a predictable fashion (i.e., according to the principle of inverse effectiveness), because the neuron's state at the beginning of the experiment was not the same at the end.

Neurons adapted rapidly in experiments where stimuli were randomly interleaved. **Figure 1** gives examples of how responses in two neurons changed over the course of an experiment. As a neuron is given repeated exposure to a particular stimulus that initially elicits a weak response (**Figures 1A,B**), the response strengthens, even when these trials are interleaved with the presentation of other stimuli. On the other hand, when a neuron is given repeated exposure to a stimulus that initially elicits a strong response (**Figures 1C,D**), the response weakens. There is some random fluctuation in the actual response magnitude on a trial-by-trial basis, but the overall trend remains consistent and is exposed by the slope of the linear regression.

The population was examined to determine if these changes were consistent. **Figure 2** shows the results of the primary analysis, in which the change (slope) of the response magnitude over trials is compared to the initial, averaged, and final response magnitudes. The first plot reveals the inverse correlation between the initial response magnitude and the slope of the response magnitude trend line. This is predicted by our fundamental hypothesis, but would also be predicted by random fluctuations: responses that were randomly large at the beginning would be expected to be weaker at the end, while responses that were randomly small would be expected to be larger at the end (regression to the mean). However, if the cause of the relationship was simple randomness, then the correlation would not improve with the averaging of multiple initial trials, which did occur, as shown in **Figure 2B** (trials 1–3; from $r^2 = 0.37$ to $r^2 = 0.39$). The average number of impulses over the entire experiment for a given stimulus condition was poorly correlated with the trend slope ($r^2 = 0.09$), as shown in **Figure 2C** (trials 1–30). Finally, there was almost no correlation between the responses on the last trial and the overall response magnitude trend ($r^2 = 0.005$; **Figure 2D**).

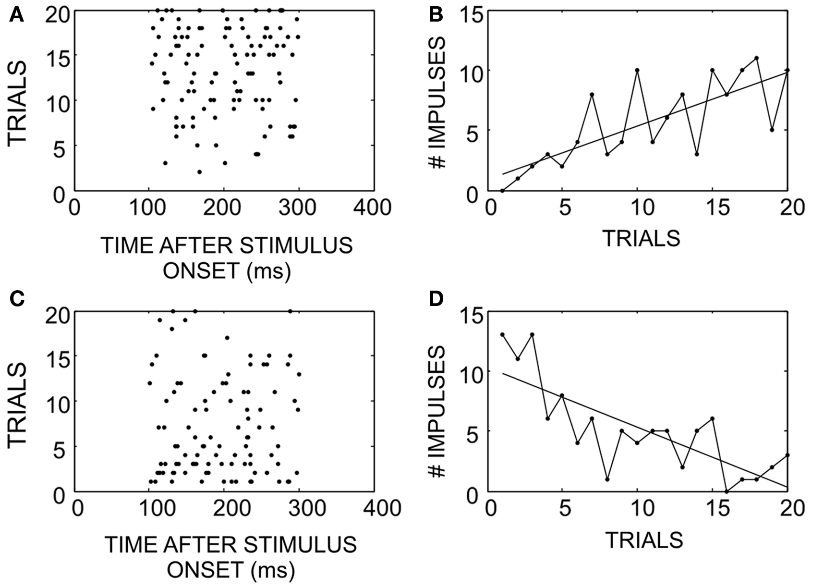

**FIGURE 1 | How SC neuronal responses change as a consequence of exposure to randomly interleaved stimuli. (A)** An example of a response that begins weak, and potentiates over successive trials (ordered bottom-to-top). **(B)** Response magnitude changes on successive trials while there is some amount of noise, the overall trend increased the number of impulses elicited. **(C)** An example of a response that begins strong, but habituates over time. **(D)** Illustration of the response magnitude changes: negative, despite random trial-by-trial noise.

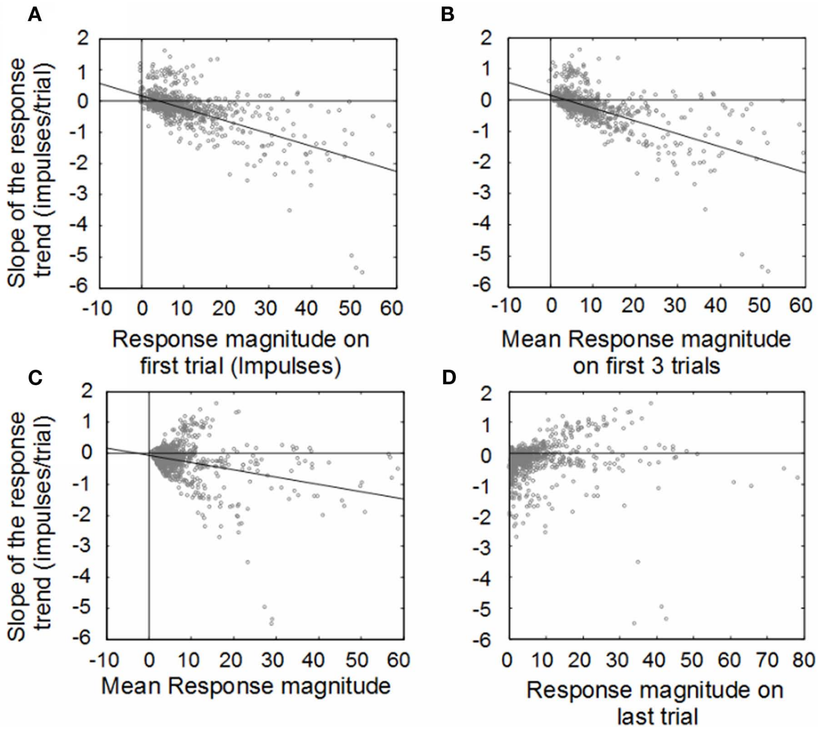

**FIGURE 2 | Illustration of correlations between the change (slope) of the response magnitude (y-axis) vs. response magnitudes (x-axis) averaged in different ways.** Change is measured by fitting a linear regression to the plot of response magnitude vs. trial number for each neuron. Both the impulse count on the first trial (far left) and the impulse count on the last trial (far right) predict the change (slope) in the response magnitude over the course of the experiment. **(A)** Weak responses observed on early trials during the experiment tended to potentiate in later trials, while strong responses in early trials tended to habituate in later trials. Averaging the response magnitudes of the first three **(B)**, trials tends to be just a slightly better predictor of the direction and magnitude of the response trend. By comparison, averaging the responses over all the trials **(C)** is a worse predictor of the direction and magnitude of the response trend, as is examining only the last trial response **(D)**.

Further support for our hypothesis is shown in **Figure 3**, where the response to a particular stimulus is correlated with the response on its next presentation (2), the presentation after that (3), and the following presentation (4). Note that there are many stimuli presented between each of these presentations. However, if our hypothesis is correct, then we expect to see that, the further away the trial in time, the less predictable the response magnitude will be. On the other hand, if our observations could be explained by simple randomness, then all responses would be equally predictive of each other (i.e., not at all). This was not the case: the more proximate the responses were in time, the better they could be used to predict one another: one-step correlations averaged $r^2 = 0.22$, two-step correlations averaged $r^2 = 0.18$, and three-step correlations averaged only $r^2 = 0.16$.

In our final analysis we compared the change in the multisensory responses over time to the changes in the best unisensory responses, which is our traditional method of determining the efficacy of multisensory integration (**Figure 4**). If the observed response changes were simply random, there would be no correlation in these slopes: the unisensory response might go up while the multisensory response might go down. Instead, we found results consistent with our hypothesis: there was a good correlation ($r^2 = 0.75$). These changes were consistent with the principle of inverse effectiveness, and further solidify the conclusion that the neuron studied at the beginning of the experiment has changed state, and so has the benefit of multisensory integration, by its end.

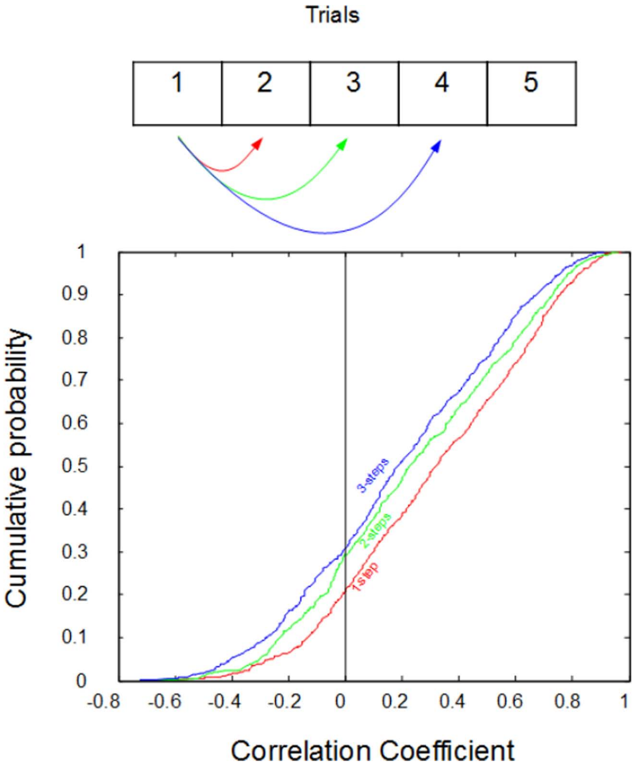

**FIGURE 3 | Cumulative distributions of the correlations between the impulse counts on subsequent trials.** Despite being randomly interleaved with other types of stimuli, the response to a stimulus on a given trial can typically be predicted from the response on its last exposure (red, "one-step"). However, predictions are progressively worse for responses from two previous exposures (green, "two-steps"), and three previous exposures (blue, "three-steps"). *r* values are plotted on the *x*-axis while cumulative probability plotted on the *y*-axis.

## DISCUSSION

The Heisenberg uncertainty principle states that certain physical properties, such as position and momentum, cannot be precisely known at a quantum level at the same time because in an effort to study one, one must disrupt the study of another (Wheeler and Zurek, 1983). We know that the nervous system is adaptive, changing its responses to stimuli to which it is exposed, that this plasticity extends well beyond the neonatal period, and that it exists at multiple levels of the neuraxis, including the SC (Yu et al., 2009). The unfortunate consequence is that we cannot study the brain without changing it, though we often take steps to attenuate these changes, for example by increasing interstimulus intervals and presenting different stimuli in a randomly interleaved fashion. The efficacy of these measures is not well-known, but has a great impact on our interpretation of the data that we collect. Here we asked a simple question: in a standard study of multisensory SC neurons exposed to randomly interleaved visual, auditory, and visual–auditory stimulus combinations, did the responses change in a predictable way? The answer was yes, and the consequences are that multisensory integration is different at the beginning of the experiment and at the end, because strong responses habituate and weak responses potentiate.

How well this principle applies to other areas of the brain, especially multisensory areas, is difficult to predict. The SC is engaged in stimulus localization and orientation behaviors (Sparks, 1986; Glimcher and Sparks, 1992; Stein, 1998), and as such, can be thought of as having two goals: the detection of salient (but potentially weak) signals, and the ignoring of events that are not novel

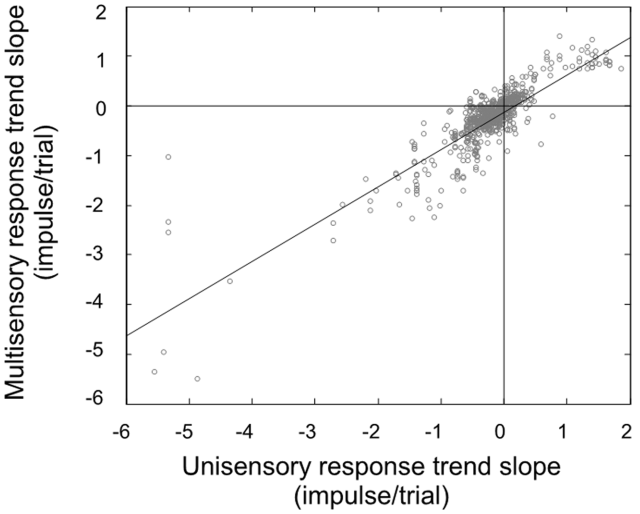

**FIGURE 4 | There is a very good correlation between the response trend slopes for multisensory and unisensory responses.** Knowing how the unisensory response will change over the course of the experiment is a good predictor of how the multisensory response will change (and vice versa).

(but potentially strong). Our observations suggest how these goals may be balanced. In circumstances in which a relevant event occurs again and again, it would be adaptive to boost the signal when it is weak. In other circumstances, where an irrelevant event repeatedly occurs, it would be adaptive to suppress its signal when it is strong. This plasticity in the neuronal response may be a general feature of the nervous system, and observable in different brain regions with different functions. On the other hand, because other areas of the brain have different functions, they may show different patterns of response changes.

That multisensory and unisensory responses appear to remain correlated throughout these changes is not surprising, and is consistent with the previously described principle of inverse effectiveness (see Stein and Meredith, 1993; Stein et al., 2009). The implications, however, are significant. When we seek to characterize the magnitude of a multisensory interaction, we must take care to appreciate not only how the dependent measure is taken, but when it is taken. Just as the impact of multisensory interactions is greatest at the onset of a response because the magnitude of the response is at its weakest (Rowland and Stein, 2007), multisensory interactions may be bigger at the beginning or end of an experiment, depending on how the neural circuit changes. This means that the neuron might be characterized one way at 2 o'clock, but at 3 o'clock look fundamentally different. Averaging data across an entire experiment may be an issue that requires additional considerations, especially with regard to statistical analyses. However, our findings suggest that a choice must be made between ignoring these changes or embracing them.

## ACKNOWLEDGMENT

Research was supported by NIH grants EY016716 and NS036916.

## REFERENCES

Alvarado, J. C., Rowland, B. A., Stanford, T. R., and Stein, B. E. (2008). A neural network model of multisensory integration also accounts for unisensory integration in superior colliculus. *Brain Res.* 1242, 13–23.

Alvarado, J. C., Stanford, T. R., Rowland, B. A., Vaughan, J. W., and Stein, B. E. (2009). Multisensory integration in the superior colliculus requires synergy among corticocollicular inputs. *J. Neurosci.* 29, 6580–6592.

Bell, A. H., Meredith, M. A., Van Opstal, A. J., and Munoz, D. P. (2005). Crossmodal integration in the primate superior colliculus underlying the preparation and initiation of saccadic eye movements. *J. Neurophysiol.* 93, 3659–3673.

Burnett, L. R., Stein, B. E., Chaponis, D., and Wallace, M. T. (2004). Superior colliculus lesions preferentially disrupt multisensory orientation. *Neuroscience* 124, 535–547.

Frens, M. A., and Van Opstal, A. J. (1998). Visual-auditory interactions modulate saccade-related activity in monkey superior colliculus. *Brain Res. Bull.* 46, 211–224.

Glimcher, P. W., and Sparks, D. L. (1992). Movement selection in advance of action in the superior colliculus. *Nature* 355, 542–545.

Jay, M. F., and Sparks, D. L. (1987). Sensorimotor integration in the primate superior colliculus. I. Motor convergence. *J. Neurophysiol.* 57, 22–34.

Jiang, W., Jiang, H., and Stein, B. E. (2002). Two corticotectal areas facilitate multisensory orientation behavior. *J. Cogn. Neurosci.* 14, 1240–1255.

King, A. J., and Palmer, A. R. (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Exp. Brain Res.* 60, 492–500.

Lee, C., Rohrer, W. H., and Sparks, D. L. (1988). Population coding of saccadic eye movements by neurons in the superior colliculus. *Nature* 332, 357–360.

McHaffie, J. G., and Stein, B. E. (1983). A chronic headholder minimizing facial obstructions. *Brain Res. Bull.* 10, 859–860.

Meredith, M. A., Nemitz, J. W., and Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *J. Neurosci.* 7, 3215–3229.

Meredith, M. A., and Stein, B. E. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science* 221, 389–391.

Meredith, M. A., and Stein, B. E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *J. Neurophysiol.* 56, 640–662.

Meredith, M. A., Wallace, M. T., and Stein, B. E. (1992). Visual, auditory and somatosensory convergence in output neurons of the cat superior colliculus: multisensory properties of the tecto-reticulo- spinal projection. *Exp. Brain Res.* 88, 181–186.

Peck, C. K., Baro, J. A., and Warder, S. M. (1993). Sensory integration in the deep layers of superior colliculus. *Prog. Brain Res.* 95, 91–102.

Perrault, T. J. Jr., Vaughan, J. W., Stein, B. E., and Wallace, M. T. (2003). Neuron-specific response characteristics predict the magnitude of multisensory integration. *J. Neurophysiol.* 90, 4022–4026.

Populin, L. C., and Yin, T. C. (2002). Bimodal interactions in the superior colliculus of the behaving cat. *J. Neurosci.* 22, 2826–2834.

Rowland, B. A., Quessy, S., Stanford, T. R., and Stein, B. E. (2007). Multisensory integration shortens physiological response latencies. *J. Neurosci.* 27, 5879–5884.

Rowland, B. A., and Stein, B. E. (2007). Multisensory integration produces an initial response enhancement. *Front. Integr. Neurosci.* 1, 4. doi: 10.3389/neuro.07/004.2007

Sparks, D. L. (1986). Translation of sensory signals into commands for control of saccadic eye movements: role of primate superior colliculus. *Physiol. Rev.* 66, 118–171.

Stein, B. E. (1998). Neural mechanisms for synthesizing sensory information and producing adaptive behaviors. *Exp. Brain Res.* 123, 124–135.

Stein, B. E., and Arigbede, M. O. (1972). Unimodal and multimodal response properties of neurons in the cat's superior colliculus. *Exp. Neurol.* 36, 179–196.

Stein, B. E., and Dixon, J. P. (1979). Properties of superior colliculus neurons in the golden hamster. *J. Comp. Neurol.* 183, 269–284.

Stein, B. E., and Meredith, M. A. (1993). *The Merging of the Senses.* Cambridge, MA: MIT Press.

Stein, B. E., and Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nat. Rev. Neurosci.* 9, 255–266.

Stein, B. E., Stanford, T. R., Ramachandran, R., Perrault, T. J. Jr., and Rowland, B. A. (2009). Challenges in quantifying multisensory integration: alternative criteria, models, and inverse effectiveness. *Exp. Brain Res.* 198, 113–126.

Wallace, M. T., Meredith, M. A., and Stein, B. E. (1993). Converging influences from visual, auditory, and somatosensory cortices onto output neurons of the superior colliculus. *J. Neurophysiol.* 69, 1797–1809.

Wallace, M. T., and Stein, B. E. (2001). Sensory and multisensory responses in the newborn monkey superior colliculus. *J. Neurosci.* 21, 8886–8894.

Wallace, M. T., Wilkinson, L. K., and Stein, B. E. (1996). Representation and integration of multiple sensory inputs in primate superior colliculus. *J. Neurophysiol.* 76, 1246–1266.

Wheeler, J. A., and Zurek, W. H. (1983). *Quantum Theory and Measurement.* Princeton, NJ: Princeton University Press.

Yu, L., Stein, B. E., and Rowland, B. A. (2009). Adult plasticity in multisensory neurons: short-term experience-dependent changes in the superior colliculus. *J. Neurosci.* 29, 15910–15922.

Zahar, Y., Reches, A., and Gutfreund, Y. (2009). Multisensory enhancement in the optic tectum of the barn owl: spike count and spike timing. *J. Neurophysiol.* 101, 2380–2394.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 29 January 2011; accepted: 16 June 2011; published online: 04 July 2011.*

*Citation: Perrault T Jr, Stein BE and Rowland BA (2011) Non-stationarity in multisensory neurons in the superior colliculus. Front. Psychology **2**:144. doi:10.3389/fpsyg.2011.00144*

*This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.*

frontiers in
PSYCHOLOGY

REVIEW ARTICLE
published: 02 May 2011
doi: 10.3389/fpsyg.2011.00077

# Organization, maturation, and plasticity of multisensory integration: insights from computational modeling studies

**Cristiano Cuppini*, Elisa Magosso and Mauro Ursino**

*Department of Electronics, Computer Science and Systems, University of Bologna, Bologna, Italy*

**Edited by:**
*Nadia Bolognini, University of Milano - Bicocca, Italy*
*Angelo Maravita, University of Milano-Bicocca, Italy*

**Reviewed by:**
*Emiliano Macaluso, Fondazione Santa Lucia, Italy*
*Alessandro Farne, Institut Nationale de la Sante et de la Recherche Medicale, France*

***Correspondence:**
*Cristiano Cuppini, Department of Electronics, Computer Science and Systems, University of Bologna, Viale Risorgimento, 2, I-40136 Bologna, Italy.*
*e-mail: cristiano.cuppini@unibo.it*

In this paper, we present two neural network models – devoted to two specific and widely investigated aspects of multisensory integration – in order to evidence the potentialities of computational models to gain insight into the neural mechanisms underlying organization, development, and plasticity of multisensory integration in the brain. The first model considers visual–auditory interaction in a midbrain structure named superior colliculus (SC). The model is able to reproduce and explain the main physiological features of multisensory integration in SC neurons and to describe how SC integrative capability – not present at birth – develops gradually during postnatal life depending on sensory experience with cross-modal stimuli. The second model tackles the problem of how tactile stimuli on a body part and visual (or auditory) stimuli close to the same body part are integrated in multimodal parietal neurons to form the perception of peripersonal (i.e., near) space. The model investigates how the extension of peripersonal space – where multimodal integration occurs – may be modified by experience such as use of a tool to interact with the far space. The utility of the modeling approach relies on several aspects: (i) The two models, although devoted to different problems and simulating different brain regions, share some common mechanisms (lateral inhibition and excitation, non-linear neuron characteristics, recurrent connections, competition, Hebbian rules of potentiation and depression) that may govern more generally the fusion of senses in the brain, and the learning and plasticity of multisensory integration. (ii) The models may help interpretation of behavioral and psychophysical responses in terms of neural activity and synaptic connections. (iii) The models can make testable predictions that can help guiding future experiments in order to validate, reject, or modify the main assumptions.

**Keywords: neural network modeling, multimodal neurons, superior colliculus, peripersonal space, neural mechanisms, learning and plasticity, behavior**

## GENERAL INTRODUCTION

The brain must deal with a complex environment where objects and events often convey a rich flow of information that simultaneously impinge to most of our senses. It is well known that information from different sensory channels is combined and integrated in the nervous system to come up with a robust and unified perception of the external world, and to provide subjects with considerable response flexibility (Stein and Meredith, 1993; Ernst and Bülthoff, 2004).

The study of multisensory integration is based on different and complementary methodological approaches, as it is exemplarily evidenced by this special issue. Neurophysiological research on animals investigates the properties of multimodal neurons in specific cortical and subcortical areas and sheds light on the basic principles that govern multisensory integration at a single neuron level (Graziano et al., 1997; Kadunce et al., 1997; Perrault et al., 2005). Experimental psychology and psychophysics characterize multisensory processes at a behavioral level, comparing response performances in tasks involving multiple modalities with respect to unimodal tasks (Driver and Spence, 1998; Farné and Làdavas, 2002; Frassinetti et al., 2002; Haggard et al., 2007). Electroencephalographic measures (such as ERP, event-related potentials) and imaging techniques (fMRI, PET) allow inferences to be drawn on the cerebral structures and neural mechanisms engaged in multisensory processes (Macaluso et al., 2000; Calvert, 2001; Eimer and Van Velzen, 2002). The previous techniques are applied both to neurologically healthy subjects and to patients with various types of sensory, attentive, and spatial disorders (Farné and Làdavas, 2002; Frassinetti et al., 2005; Sarri et al., 2006) – that may differently affect multisensory abilities – to gain further insight into the neural correlates of multisensory integration.

The previous approaches have provided a great body of data on the topic, and have contributed to characterize properties of multisensory integration and identify the cerebral areas mainly implicated in this phenomenon. However, the comprehension of the neural mechanisms by which this brain capability is realized is still insufficient. This limitation may in part be ascribed to the complexity of the mechanisms involved; indeed, multisensory integration plausibly arises as a emergent property of interconnected neural populations, in which many factors such as the characteristics of the single neurons, arrangement of the connections, network topology, integrity or impairment of some circuits contribute to determine the observed effects. Clarifying these aspects is quite arduous based on experimental results only. Moreover, the lack of an adequate knowledge on the neural topology and connections underlying multisensory integration significantly limits the comprehension of the neural learning mechanisms through which multisensory integration capabilities are acquired. Indeed, many data in the literature indicate that the ability to integrate sensory information is not innate in the nervous system, rather it gradually develops and may plastically change with sensory experience; that is, the experience with the external world, rich of cross-modal stimuli, would shape network in a functionally relevant manner. The learning rules and the conditions that drive maturation and plasticity of multisensory integration in the brain are still far from being well understood.

In order to improve understanding of computational principles and neural mechanisms of multisensory integration, in recent years the traditional research approaches have been assisted by the use of computational models and digital simulation techniques. The proposed models can be roughly divided into two main categories: Bayesian models and connectionist models.

Bayesian models consider the problem of sensory cue integration within the theory of statistical inference (Anastasio et al., 2000; Colonius and Diederich, 2004). They provide a mathematical framework within which multisensory effects (both at behavioral and at neuronal level) can be accounted for, but they do not gain insight into how the necessary computation is neurobiologically performed.

Connectionist models make use of artificial neural networks, and are particularly suitable to formalize hypotheses on the learning mechanisms and neural circuitry underlying multisensory integration. This type of models emulate some fundamental characteristics of the biological neural networks, that appear to have a key role in multisensory integration: the collective behavior of the interconnected neurons gives rise to emergent properties that are not possessed by the single network components; moreover, the network may learn from its inputs and shape its behavior, by modifying the weights of its synaptic connections. A number of these models have been proposed in the literature (Pouget et al., 2002; Anastasio and Patton, 2003; Avillac et al., 2005; Martin et al., 2009) focused on different aspects of multisensory interactions and tied on specific multisensory neural regions.

In this paper, we present two neural network models of multisensory integration, that we recently developed. The two models tackle two different and specific problems, that have been received growing attention in the last decades within the multisensory research community, and for which a great body of data have been collected.

The first model (Magosso et al., 2008; Ursino et al., 2009; Cuppini et al., 2010) considers the integration of visual and auditory stimuli, as it occurs in the superior colliculus (SC), a midbrain structure implicated in driving overt responses (such as eyes and head movements) toward external events. The deep layers of SC are a robust locus for multisensory integration and have provided a fertile site in which to examine this phenomenon. The proposed model is able to emulate the features of multisensory interaction experimentally observed in SC neurons, and to explain how the development of these abilities may be guided by sensory experience.

The second model (Magosso et al., 2010a,b) treats the problem of how visual stimuli or auditory stimuli close to the body (for instance stimuli on and close to the hands) interact with tactile stimuli to form the perception of peripersonal space (i.e., the space immediately surrounding our body). The model identifies network architecture and connections able to reproduce several data on multisensory representation of peripersonal space, and hypothesizes some physiological mechanisms to account for the plastic changes of peripersonal representation as a function of experience.

In the following, for each model we will describe the physiological counterpart, the model structure, and simulation results. The emphasis will not be on mathematical details and on implementation of the model. Rather, by considering these two exemplary cases of multisensory integration, we aspire to evidence the potentialities of computational models to gain insight into the neural mechanisms underlying organization, development, and plasticity of multisensory integration in the brain. In particular, we will show how by using mathematical models plausible scenarios can be formalized in quantitative terms and knowledge obtained using different approaches can be synthesized into a unique, coherent structure; how models may help the interpretation of behavioral and psychophysical responses in terms of the reciprocal interconnections among neurons; how, neural network modeling may be integrated with experimental research, by generating new predictions and suggesting novel experiments, to promote progress in the comprehension of multisensory integration processes.

## AUDIO–VISUAL INTEGRATION IN SUPERIOR COLLICULUS: A NEURAL NETWORK MODEL

### BACKGROUND

Let us consider the problem of integration of visual and auditory stimuli to drive overt behavior. The concepts described below refer to a particular midbrain area, the SC, which has been deeply studied in the context of multisensory integration: however, they may have a more general validity and are suitable to illustrate how a biologically inspired neural network can realize multisensory integration to improve the response to external stimuli.

The role of the SC is to initiate and control overt movements in response to important stimuli from the external world, for instance to control the shift of gaze or to orient various sensory organs to a correct direction (Stein and Meredith, 1993). It receives stimuli from various brain regions involved in auditory, somatosensory, and visual processing (Edwards et al., 1979; Huerta and Harting, 1984; Stein and Meredith, 1993).

While some neurons in the SC are unisensory, more than half are multisensory, i.e., they respond to stimuli of different sensory modalities. Multisensory neurons in general have receptive fields (RFs) for different modalities in spatial register; this means not only that a visual–auditory neuron will have two RFs (one for the auditory and one for the visual modality) but these RFs have a large superimposed region (Meredith and Stein, 1996). These RFs are topographically organized, so that proximal neurons in the SC have RFs with proximal centers in the environment.

The presence of multisensory neurons, whose RFs are in spatial register, can explain a phenomenon named "multisensory enhancement:" when two cross-modal stimuli (for instance one visual and one auditory) come from proximal positions of space and in close temporal proximity, the response of the SC neuron is generally greater than each of the individual unisensory responses (Kadunce et al., 2001; Perrault et al., 2005). Furthermore, the response of a multisensory SC neuron follows a rule named "*inverse effectiveness*:"

the enhancement produced by two spatially aligned cross-modal stimuli is inversely related to the effectiveness of the individual modality-specific components (Perrault et al., 2005).

The complexity of the SC response, however, is much greater than that emerging from a single non-linearity, i.e., from the behavior of a single neuron. Several other aspects, related with the interactions among neurons should be considered.

First, if two within modal stimuli (i.e., two stimuli of the same modality, for instance both auditory or both visual) or two cross-modal stimuli (i.e., stimuli of different modalities, one auditory and the other visual) originate from disparate positions in space, the final response of the SC can be reduced or eliminated compared with the response to an individual stimulus alone ("*within modal and cross-modal suppression*;" Kadunce et al., 1997). This behavior implicates the presence of some competitive interactions among neurons whose RFs are located at different spatial positions.

Finally, several experimental data were collected recently to analyze how these multisensory neurons in the SC acquire integrative capabilities. After few weeks from birth many SC neurons are multisensory (i.e., they respond to inputs of different sensory modalities) but are not able to integrate them. The integrative capability appears only after several weeks and after a protracted cross-modal experience (Wallace and Stein, 1997; Wallace et al., 2004).

A further important aspect, which seems strictly related with the maturation of multisensory integration, concerns the input pathways which converge to the SC: these include both ascending pathways from subcortical zones and descending inputs from the cortex (mainly from a region named the anterior ectosylvian sulcus (AES); the latter, in turn, includes a visual area, AEV, and an auditory area, FAES). Stein et al. (Wallace and Stein, 1994; Jiang et al., 2001; Alvarado et al., 2009) demonstrated that the capacity to integrate multisensory inputs (either enhancement or depression) depends on the presence of an intact cortex. If cortical inputs to the SC are entirely or selectively removed, SC neurons remain multisensory (although with a reduced response) but lose their integrative capacity.

Some authors formulated the hypothesis that maturation of multisensory integration in the SC strongly depends on the formation of descending synapses from the cortex (Wallace et al., 1993; Wallace and Stein, 2000; Jiang et al., 2006, 2007). In the kitten, only ascending inputs would be effective, although weak and with a poor spatial resolution. Descending synapses would maturate under pressure of a cross-modal environment, to store the statistics of multisensory events occurring early in life, in order to optimize the probability of a correct event detection.

The analysis of neural mechanisms involved in multisensory integration, both in early life and after maturation, is not only important for physiology, in order to gain a deeper comprehension of how the SC realizes its function, but may also help understanding complex behavioral responses in humans. In this regard, a model of the SC that summarizes the main experimental findings and elucidates possible mechanisms, may represent a good starting point for understanding the common role of multisensory integration in overt behavior.

Previous important models were especially focused on information theory. In particular, Anastasio, Patton et al. (Anastasio et al., 2000; Patton et al., 2002; Patton and Anastasio, 2003) developed models in which the SC neurons implement the Bayes rule to compute the conditional probability that a target is present in their RF. These models were able to reproduce cross-modal enhancement as well as within-modal suppression but were not inspired by neurobiological mechanisms. A similar approach was used by Colonius and Diederich (2004) by using the maximum likelihood. By modeling a network of the corticotectal system and using a learning algorithm, Anastasio and Patton (2003) were able to simulate self-organization in the corticotectal system with the formation of neurons with and without multisensory enhancement. However, their model neglects the important fact that different circuit components appear to play different roles in multisensory integration. A single-neuron model was proposed by Rowland et al. (2007). The model shows results which resemble empirical findings (multisensory enhancement, superadditivity, inverse effectiveness, the effect of NMDA-receptor deactivation, and temporal disparity); however the model does not incorporate the fact that the individual SC neuron is embedded in a network in which interactions between units can affect responses.

In previous years we presented a model (Magosso et al., 2008; Ursino et al., 2009; Cuppini et al., 2010), which is inspired by biological mechanisms and can explain most of the results delineated above. Furthermore, a last version of the model explains the maturation of the SC integrative capabilities.

In the following, the main aspects of the model are first presented and justified. Then, some simulation examples are shown and commented on the basis of the mechanisms incorporated in the model. In the last section, model implications for learning and for behavior are stressed, thinking to a more general perspective.

## MODEL DESCRIPTION

A qualitative sketch of the model is given in **Figure 1**. Fundamental aspects are explained below while all equations, mathematical details and parameter numerical values can be found in previous publications of the authors (Magosso et al., 2008; Ursino et al., 2009; Magosso et al., 2010a).

- Each neuron is described through a sigmoidal relationship (with lower threshold and upper saturation) and a low-pass filter (which simulates the dynamics of the neuron, i.e., the time required to reach a steady-state condition in response to a sudden input change). Neurons normally are in a silent state (or exhibit just a mild basal activity) and can be activated if stimulated by a sufficiently strong input. *In vivo* the sigmoidal non-linearity can be ascribed to the typical characteristics of neurons, which need a sufficient input current to generate spikes and which saturate: this behavior may be further accentuated by non-linearities in the receptor responses at the synapse levels (for instance, the response of NMDA receptors). Low-pass dynamic can be ascribed to the response of the cell membrane and to the synaptic response.
- The model is composed of four unisensory areas (see **Figure 1**). Two represent the visual and auditory subregions of the AES cortex which send descending pathways to the SC (respectively AEV area and FAES area); the other two areas are responsible for all other (ascending) visual and auditory input sources (non-AEV and non-FAES areas). These four input

regions respond only to modality-specific inputs: AEV and non-AEV are sensitive to visual stimuli, while FAES and non-FAES to auditory ones. This arrangement has been chosen to reproduce the importance of AES inputs in driving the SC responses, with respect to all other input sources. For simplicity elements of each area are organized in a one-dimensional chain, and preserve a topological organization, i.e., proximal neurons respond to stimuli in proximal position of space.

- Each element of the unisensory areas has its own RF that can be partially superimposed on that of the other elements of the same area. The elements of the same unisensory area interact via lateral synapses, which can be both excitatory and inhibitory. These synapses are arranged according to a Mexican hat disposition (i.e., reciprocal excitation among neighbors and reciprocal inhibition with distant elements).
- The model also includes four different populations of inhibitory interneurons. Each interneuron receives stimuli from just one unisensory area (hence we have four distinct interneuron populations, see **Figure 1**) and works to inhibit some inputs to the SC. In particular, interneurons which receive their inputs from non-FAES and non-AEV areas realize a competitive mechanism between the two ascending pathways, so that only the stronger ascending input may affect the SC. The interneurons which receive their inputs from the AES (i.e., from the descending pathway) inhibit ascending inputs to the SC. Hence, in the presence of descending inputs, the ascending inputs are ineffective.
- Finally, a multisensory area represents neurons in the SC responsible for cross-modal integration. The elements of this region receive inputs from neurons in the unisensory areas (AES and non-AES unisensory regions) and from the interneuron populations. Moreover, elements in the SC are reciprocally connected by lateral inhibitory or excitatory synapses with a Mexican hat disposition.

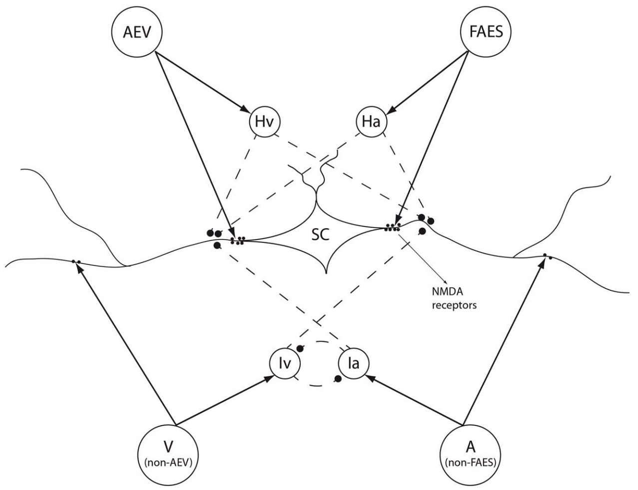

**FIGURE 1 | The general structure of the superior colliculus (SC) network.** The four projection areas make excitatory synapses with their target SC neurons and with their target interneurons (solid black arrows). The interneurons, by means of inhibitory synapses (dashed black lines), provide two competitive mechanisms: (1) Ha and Hv provide the bases through which the inhibitory effect of AES is imposed on non-AES inputs; (2) Ia and Iv provide the substrate for a competition between two non-AES inputs in which the stronger one overwhelms the weaker.

## RESULTS

In the following, we separately present the integrative behavior reproduced by the model and we discuss how the different aspects of the model contribute to explain the main results on multisensory integration. In particular, we analyze the multisensory integrative abilities of the SC (cross-modal enhancement and depression), and the role played by the AES cortex in eliciting these phenomena. In a final section, we analyze how these capabilities are acquired during postnatal maturation depending on sensory experience with cross-modal events.

### *Multisensory integration*

#### *Cross-modal enhancement*

Results in the literature suggest that the response of SC neurons to cross-modal stimuli in spatial register is greater than the response to any individual unisensory stimulus (a phenomenon named as enhancement). However, measured in percentage of the stronger unisensory response, the enhancement is greater when the individual stimuli are weak, otherwise known as the "principle of inverse effectiveness" (Meredith and Stein, 1986; Stein and Meredith, 1993; Wallace et al., 1998; Perrault et al., 2003, 2005; Stanford et al., 2005; Stein et al., 2009). These observations are common among SC neurons.

To reproduce this phenomenon, we stimulated the network with two modality-specific stimuli (one auditory and one visual) located at approximately the same position in space. These inputs are presented both simultaneously (cross-modal configuration) and independently (modality-specific presentation), at different levels of efficacy.

As illustrated in **Figure 2**, the model accounts for the main results reported in the empirical literature: (a) the model produces multisensory enhancement for each level of input stimuli;

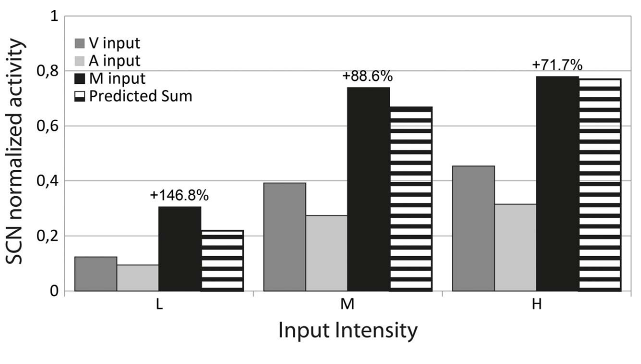

**FIGURE 2 | Multisensory Enhancement and Inverse Effectiveness in the model.** Activities evoked in the SC neurons in response to Visual (V input, dark gray bars), Auditory (A input, light gray bars) and Cross-modal (M input, black bars) stimuli at different levels of efficacy, placed at the center of the RF. The intensity of the stimuli is plotted in the *x*-axis: L, low efficacy input; M, medium efficacy input; and H, high efficacy input. For each level we display the percent enhancement produced by the cross-modal configuration, and the predicted sum (striped bars), that is the sum of the responses evoked by the single modality-specific components of the multisensory stimulus.

(b), enhancement is greater (about 150%) when small stimuli are used as input, and decreases (about 70%) when strong inputs are used, in agreement with the principle of inverse effectiveness; (c) the model shifts from a superadditive computation to an additive computation at higher levels of stimulus effectiveness, and (d) auditory stimuli are less effective than visual stimuli to elicit the SC response.

The previous results can be explained by the following characteristics of our model: (i) the presence of unisensory areas, with modality-specific RFs; (ii) the presence of a multisensory area, whose neurons have auditory and visual RFs in spatial register, (iii) the presence of a sigmoidal relationship for neurons. A small modality-specific input cannot be strong enough to produce a significant response in the sigmoidal function of the SC neuron, but if it is coupled with another weak stimulus, this combination could produce an appreciable result in the sigmoidal curve. This explains the strong percentage enhancement evident with weak inputs. Conversely, if the input are strong, two cross-modal stimuli lead the SC neuron close to saturation, thus resulting in a reduced enhancement.

### *Modality-specific and cross-modal suppression*

Several experimental results (summarized in the introduction) reveal that a second spatially distant (cross-modal or modality-specific) stimulus, causes depression in the response of the SC neuron to a first stimulus located inside its RF. This means that distal stimuli induce a competition among SC neurons. To explain cross-modal suppression, we assumed the presence of lateral synapses among multisensory neurons in the SC, with a Mexican hat disposition: proximal neurons send reciprocal excitatory connections, but exchange inhibitory connections with more distal neurons. It is worth noting that this arrangement of lateral synapses can explain both cross-modal and within-modal suppression.

The dependence of cross-modal integration on the spatial configuration of the stimuli is shown in **Figure 3**. In this simulation we used a constant strong visual stimulus located at the center of the RF of the target neuron, and a second strong auditory stimulus placed at different locations in space. The simulations have been repeated by varying the distance between the two stimuli, and examining its effect on the response of the SC neurons. As far as the stimuli are in spatial proximity (i.e., both are inside the RF of the same multisensory neuron, relative distance less than 5°), the cross-modal configuration produces multisensory enhancement, in agreement with **Figure 2**; conversely, when the two modality-specific stimuli are placed far apart, the resulting activity in the SC is depressed (first two panels on the left in **Figure 3**). Depression is greater than 50%.

According to the model, a single mechanism (i.e., lateral inhibition within the multimodal area) can explain both within-modal and cross-modal suppression. However, results in the literature indicate that within-modal suppression may occur in the absence of cross-modal suppression (whereas the reverse behavior is never true, i.e., cross-modal suppression always occurs together with within modal suppression). According, in our model within-modal suppression is affected not only by the presence of lateral inhibition within the multisensory area, but also by lateral synapses arranged as a Mexican hat operating at the level of the unisensory areas. If the suppressive mechanism in the unisensory area is weak compared with that in the multimodal area, within-modal and cross-modal suppression have approximately the same strength. Conversely, if we assume the existence of strong inhibitory synapses in one unisensory area, but poor inhibitory synapses in the SC area, we may simulate strong within-modal suppression without cross-modal suppression. Examples of the latter behavior, which has been experimentally observed in some SC neurons (Kadunce et al., 1997), are illustrated in our previous works (see Magosso et al., 2008).

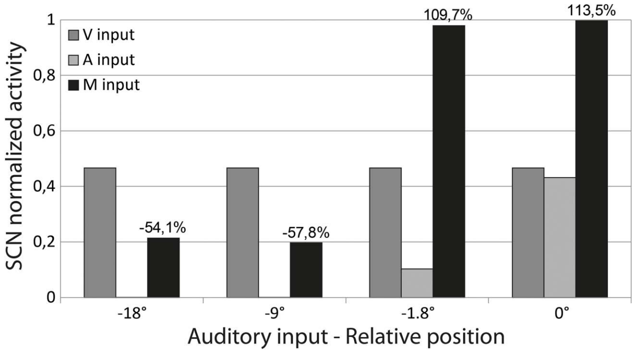

**FIGURE 3 | Multisensory integration with respect to stimuli relative position.** Simulations were performed by applying a visual stimulus of high intensity at the center of the neuron RF, and moving a second auditory stimulus (of high intensity) far from the receptive field. The responses evoked by the individual stimuli (acting separately) and the cross-modal response are reported at different distances (distance is plotted in the *x*-axis). When the two modality-specific stimuli are both inside the RF of the analyzed SC neuron, both evoke a response and, in a cross-modal configuration, produce a multisensory enhancement. Conversely, when one is inside the RF (V input here, dark gray bars), and the second is outside (A input here, light gray bars) only the first drives an activity in the SC neuron. However, the cross-modal stimulation results in a depressed activity of the observed neuron. Percent Enhancement and Depression are reported in the figure.

### *AES role*

Recent empirical data reveal that deactivation of AES eliminates multisensory integration in SC neurons, whereas it just moderately reduces their unisensory responses (Wallace and Stein, 1994; Jiang et al., 2001; Alvarado et al., 2007, 2009). The same essential observation is made when individual subregions of AES are deactivated (e.g., AEV or FAES, see Alvarado et al., 2009). However, in the latter case, only the responses that are sensitive to inputs from that region are affected (Alvarado et al., 2009).

To analyze the responses of the model in case of full and partial AES inhibition, we repeated the same simulations presented in **Figure 2**, using very effective stimuli (case H in **Figure 2**), and (i) by selectively deactivating the overall AES; (ii) by deactivating the AEV only; (iii) by deactivating the FAES only. The results are reported in **Figure 4**.

When the entire AES cortex is deactivated, the unisensory responses are smaller (reaching only ~20% of the maximum activity), a finding that parallels the physiology. Also, and more importantly, the multisensory response is not significantly greater than the response to the more effective of the two component stimuli: hence multisensory enhancement is no more present.

The same finding is evident when AEV only or FAES only are separately deactivated: even subregional deactivation eliminates multisensory enhancement. However, in this condition the effect of deactivation is modality-specific: deactivation of AEV affects the visual responses but not the auditory responses. The reverse occurs with deactivation of FAES.

These results can be explained by the presence of inhibitory mechanisms in the model. In particular, in the complete absence of AES, the two ascending inputs (from non-AEV and non-FAES areas in **Figure 1**) compete so that just the stronger input affects the target SC neuron. The competition results in a multisensory response no greater than the response to one of the component stimuli. In case of partial deactivation, the intact AES region completely suppresses all non-AES inputs through the descending interneuron populations. As a consequence, when a cross-modal stimulus is presented, the stimulus in the non-deactivated modality dominates the response.

### *Maturation of multisensory integrative capabilities*

As shown above, in the adult cats the SC presents the ability to integrate stimuli of different sensory modalities to drive an appropriate behavioral response to external events. This capability is yet not present at birth. Several experimental findings have shown that in the kittens – even after several weeks – the SC is multisensory, but not able to integrate (Wallace and Stein, 1997). Here we present some results to show how the model is able to reproduce the maturation of this structure in the first weeks after birth, assuming a given disposition of the synapses at birth and using reliable rules for synaptic plasticity.

In order to reproduce the neonatal condition, we assumed that the descending synapses from AES are just virtual, and their effect to the SC neurons is negligible. Moreover, ascending projections from non-AES regions are weak and with a widespread spatial disposition (hence, the RFs of SC neurons are very large). Under these conditions we performed the same set of simulations as in **Figures 2 and 3**, to simulate the behavior of a neonatal SC (**Figures 5A,C**). Subsequently, we simulated the maturation process by means of an Hebbian training, performed by presenting thousands of stimuli to the network, both cross-modal and modality-specific. More particularly, the training rule is based on the following points: (i) a synaptic potentiation if the pre-synaptic and the post-synaptic neurons are both active above a given threshold; (ii) synaptic depotentiation if the pre-synaptic neuron is inhibited while the post-synaptic neuron is active above a given threshold; (iii)normalization of synapses, so that the sum of synapses entering a neuron does not overcome a given maximum saturation value. All these aspects are physiologically reliable. Finally, the same set of simulations was repeated to analyze the SC behavior after training (**Figures 5B,D**).

**Figure 5** shows the results of these simulations both in the neonatal configuration before training (on the left), and in the adult condition at the end of training (on the right). In the simulated neonate the SC is able to respond to different modality-specific stimuli, but it does not present integrative capabilities, neither

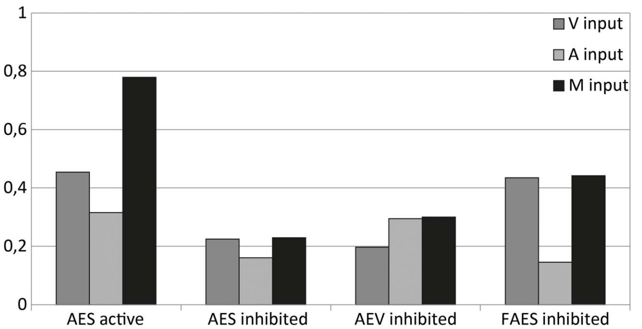

**FIGURE 4 | Behavior of the network as function of AES cortex.** These figures compare the activity of SC neurons in response to different inputs with AES active or inhibited, fully (AES inhibited) or only partially (AEV inhibited, FAES inhibited). In all simulations, the activity was assessed by stimulating the model with auditory (A input, light gray bars), visual (V input, dark gray bars), and multisensory (M input, black bars) inputs at a very high intensity (H level in **Figure 2**). If the AES is totally inhibited, the SC shows no multisensory integration, the unisensory responses are reduced by about 50% and the response to two cross-modal stimuli looks like the stronger unisensory one. If just the AEV is inhibited, the SC presents a normal response to an auditory stimulation, but the response to a modality-specific visual stimulation is reduced by about 50% compared to that produced when AEV is active. The multisensory response looks like the stronger one (in this case the auditory one). In case of FAES inhibited: the SC response to a visual stimulus is unaffected whereas the response to an auditory stimulus is depressed compared with the intact case; multisensory stimulation elicits a response similar to the visual one. The stimuli were presented in the center of the RF of the observed SC neuron. Note the loss of multisensory integration when AES is deactivated even partially. Multisensory integration capability needs both AES subregions active.

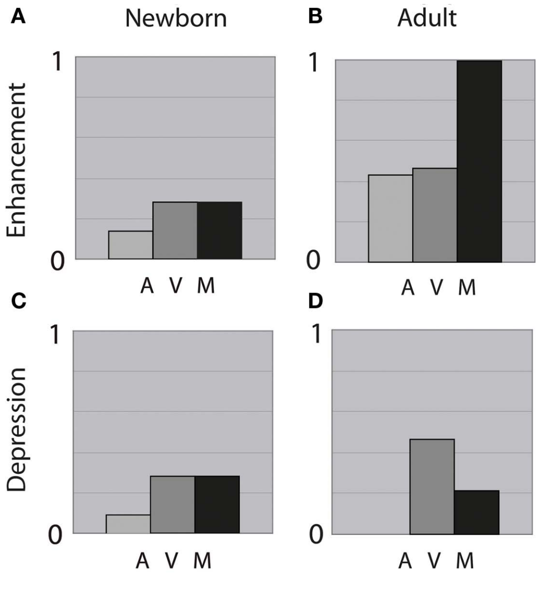

**FIGURE 5 | Integrative capabilities in the neonate and in the adult cats. (A,B)** Shows the responses obtained using two stimuli (one auditory and the other visual) of high intensity placed at the center of the RF. **(C,D)** Shows the responses evoked by a visual stimulus at the center of the RF, paralleled by a distant auditory stimulus (relative distance = 9°). In **(A,C)**, the neonatal SC neuron is incapable of integrating the cross-modal inputs and has responses equivalent to those of the stronger of the two modality-specific component. In **(B,D)**, the adult SC neuron exhibits both multisensory enhancement and depression.

enhancement nor depression. Conversely, after training, the observed SC neuron has acquired the ability to integrate stimuli of different sensory modalities, in different spatial configurations. These results cope quite well with data present in the literature (see for instance Figure 10 in Wallace and Stein, 1997).

### *Statistical analysis*

Finally, in order to compare model behavior in the three configurations (immature, adult intact, adult without AES) we performed some statistical tests. To test end, we generated 200 pairs of spatially aligned random stimuli (200 visual and 200 auditory) ranging from a value just below the threshold for the unisensory neuron to a value close to saturation (in order to exploit the overall dynamic range of neurons), with a uniform distribution. For each pair of stimuli, the SC response was computed to any unisensory stimulus, and to their cross-modal combination. This set of simulations was repeated for each configuration of the network (immature, adult intact, adult with no AES).

The results are summarized in **Figure 6** (mean + SD). Two aspects of this figure are worth noting: (i) the strong increase of the cross-modal response in the adult compared with the cross-modal response in the immature, and (ii) the disappearance of the cross-modal enhancement after AES deactivation. Finally, we compared the population of cross-modal responses in the intact adult with the populations of cross-modal responses in the immature and in no-AES cases, and with the populations of unisensory responses in the intact adult (Mann–Whitney test). All differences turned out highly significant ($p < 0.0001$).

## DISCUSSION

The model presented above is able to explain many different experimental results on multisensory integration in the cat's SC, assuming reliable mechanisms for cross-modal integration. However, although the model was built to investigate a single neural structure in a specific animal, we claim the proposed mechanisms may have a more general validity for the problem of sensory fusion, well beyond the particular physiological system considered. Hence, in this ensuing discussion, the importance of the mechanisms will be analyzed thinking to the general problem of how senses can be merged, and underlying their presumed impact for a correct overt behavior in response to multisensory events.

### *Enhancement and inverse effectiveness*

In our model, in its adult configuration, the response to two cross-modal stimuli in close spatial and temporal proximity turns out much stronger than the response to any individual unisensory stimulus. Moreover, enhancement is more evident in response to weak stimuli than to stronger ones, a behavior that can be ascribed to the presence of a sigmoidal characteristic for neurons. The impact of "inverse effectiveness" for overt behavior is evident. A weak unisensory stimulus alone may not contain enough information to drive the behavior and may be easily confused with noise or not discriminated from alternative proximal events (see also point ii below, on depression). However, the reliability of an event increases dramatically if two cross-modal stimuli occur together, a condition frequently met in our daily life. It is worth noting, however, that the last behavior is not innate, but is learned on the basis of the interaction with the external environment (see point iv below, on maturation). This idea resembles, although in different form, the idea proposed by Anastasio and Patton (Anastasio et al., 2000), according to whom SC neurons detect the conditional probability of an external event.

### *Cross-modal and within-modal suppression*

An important result, which has serious consequences on behavior, is that two distal stimuli (either cross-modal or within-modal) compete reciprocally, thus causing a depressed response. This competition is maximal at moderate distances (about 15–20°) but decreases at larger distances. Model ascribes this behavior to the presence of lateral inhibitory synapses among neurons in the same area. In particular, the present model assumes the presence of lateral synapses, with a Mexican hat arrangement, in all areas (both in the unisensory areas, AES and non-AES, and in the SC). A Mexican hat disposition is frequently assumed in the cortex, not only in modeling primary perceptual areas but also in higher associative areas (such as the parietal and frontal cortices; Amari, 1989; Mascaro et al., 2003). Hence, cortical aspects of the model (here the AES) are well motivated. Conversely, it is more difficult to find neurophysiological results which motivate a Mexican hat disposition in subcortical structures, although this kind of interaction can be found in the initial processing pathways (for instance in the retina). Hence, this disposition can be justified only "*a posteriori*" on the basis of obtained results, and may represent a testable aspect

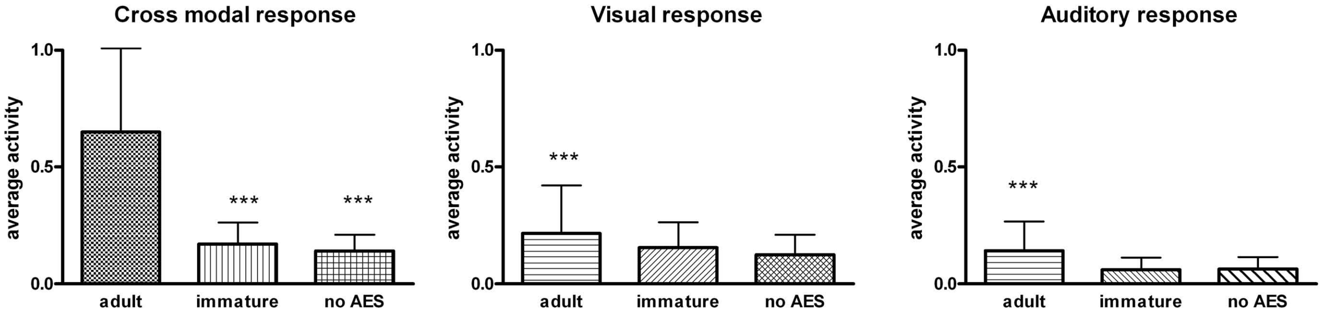

**FIGURE 6 | Model responses (mean + SD) to 200 randomly generated pairs of stimuli, in the three network configurations: immature, adult intact, adult with no AES.** The cross-modal response in the adult intact is significantly different ($p < 0.0001$) compared with the cross-modal response in the immature and in the no-AES configurations, and also compared with the unisensory responses in the intact adult (Mann–Whitney tests).

of the model. In our model lateral synapses in the SC play a pivotal role to generate cross-modal depression to misaligned stimuli in the adult. Without these synapses, cross-modal suppression would not occur. Lateral synapses in the non-AES areas have a less definite role: they produce a certain within-modal depression, which becomes evident in case of AES suppression.

### *Ascending vs. descending inputs*

An important aspect of the last model version is the different role played by ascending (subcortical) and descending (cortical) inputs to the SC, in agreement with experimental results. Although this arrangement reflects our anatomo-physiological knowledge on the SC, it may lead to interesting considerations applicable to more general sensory-fusion problems. The fundamental aspect is that the SC possesses two alternative routes to receive multisensory inputs, and these have different characteristics. Ascending inputs to the SC are able to induce a multisensory response (in our exempla, a response to both auditory and visual stimuli); however, this specific pathway does not result in any clear multisensory integration (in particular, no enhancement is evident). Only the stronger unisensory ascending input determines the final response. To simulate this behavior (which becomes evident in the adult network after deactivation of the AES, and is also evident at birth), model assumes that ascending inputs interact through a competitive mechanism. Competitive mechanisms are frequently encountered in networks which process perceptual inputs, and may help the formation of a clear-cut response excluding unnecessary inputs. Conversely, the two descending inputs (originating from AEV and FAES) induce a strong multisensory integration, that is the typical behavior of an adult and provides a better response to a multisensory environment. Furthermore, in order to reproduce experimental findings, the model assumes that the descending pathways completely inhibit the ascending ones, thus dominating the adult behavior.

An important question, at this point, is: why the SC exhibits these two alternative input paths? And what may be their specific significance for behavior? We have two possible responses to these questions. First, physiological systems always present a certain amount of redundancy: this means that certain mechanisms, normally silent, may become effective in particular exceptional conditions. In our model, the ascending inputs may assume a role in the presence of neurological deficits, for instance after a lesion of the cortical structures converging to the SC. This aspect may be of importance for the neuroclinics, and might be exploited in future works to drive rehabilitation procedures, for instance by using the ascending paths to induce synaptic plasticity. A second possible role of the ascending path is in driving maturation, as discussed in last point iv below.

### *Maturation of multisensory integration*

According to recent experimental results (Wallace et al., 1993; Wallace and Stein, 2000; Jiang et al., 2006, 2007), we assumed that the ascending route provides the dominant inputs at birth, whereas descending inputs are just latent at this stage. Moreover, the ascending synapses at birth are weak and exhibit only a poor spatial resolution. They do not code for the statistics of the external world, but simply set the SC to an initial working condition, characterized by a moderate spatial arrangement for neurons and a moderate non-integrative response to the stronger multisensory stimuli. In this schema, the ascending pathways would have the role to set an initial bias to drive learning. Conversely, the descending pathways, and the related cortical structures, would have the role of learning and storing the statistics of the external environment. If the subject experiments many cross-modal events, with visual and auditory stimuli in close spatial and temporal proximity, synapses form AEV and FAES exhibit a simultaneous Hebbian reinforcement, which is at the basis of multisensory integration. Conversely, if external stimuli are commonly unisensory, only one kind of synapse reinforces (for instance, those from AEV if we assume visual stimulation only) whereas the other ones (from AES) are reduced through Hebbian depotentiation, thus preventing the formation of multisensory integration. Thus, model predicts that multisensory integration requires the presence of concurrent cross-modal stimuli, but can be forgot if the subject is exposed to a unisensory environment. It is worth noting that also the inhibitory descending synapses are learned in our model, as well as the lateral synapses within the SC: this aspect explains the appearances of cross-modal depression among distal stimuli, and the predominance of the descending pathways on the ascending ones. After a long training process in a multisensory environment (as the one in which we live normally) the descending integrative pathways completely dominate behavior and suppress the role of the ascending path. However, the ascending paths prune their spatial resolution during training, and may replace the descending ones in case of cortical deactivation.

According to the previous analysis, we expect that the SC model, without further assumptions (or just by better assessing some parameter values) can replicate maturation in a different environment. For instance, if dark reared cats were simulated (absence of visual stimuli during the training), visual descending synapses would never be created, and SC neurons would not develop multisensory integration. In case of cross-modal inputs with spatial disparity, SC neurons in the model would receive descending synapses originating from distal positions, and so would develop multisensory integration for spatially disaligned cross-modal stimuli. Preliminary simulations (not reported in this paper) confirm these suppositions.

Finally, it is of value to underline some model limitations, and point out lines for future improvements. A limitation is that the training period was started with the same ascending synapses for all neurons. In other words, we used a deterministic pattern of initial synapses in the ascending path, and the sole random aspect consists in the nature and position of the stimuli generated during the training. We claim that wider differences among neuron behaviors at the end of the maturation, including the presence of some not-integrative neurons, may be obtained using a random disposition for the ascending synapses at the beginning of the training. This may be plausible, since ascending synapses maturate during the first 4 weeks (in the cat): after this period, they are certainly not everywhere equal.

A further limitation is that we used just a single statistics for the input stimuli during the training. It is probable that increasing the percentage of unisensory inputs would increase the number of neurons which do not develop multisensory integration after the training, due to the presence of a forgetting factor in the learning rule.

In conclusion, the present SC model may provide important suggestions on which neural mechanisms may be responsible for cross-modal enhancement and inverse effectiveness; on which mechanisms may explain response suppression in the presence of ambiguous or conflicting stimuli; and on how multisensory integration can develop under the pressure of an external environment, starting from a moderate initial spatial bias of neurons, and exploiting the statistics of the external stimuli.

## MULTISENSORY REPRESENTATION OF PERIPERSONAL SPACE: A NEURAL NETWORK MODEL

### BACKGROUND

The near space (peripersonal space) is behaviorally and functionally distinct from the far space (extrapersonal space; Rizzolatti et al., 1997) since objects within it can potentially enter in contact with our body. Depending upon their nature, near objects could be either avoided or reached and manipulated.

Evidence for a specific representation of the peripersonal space and for its properties have first come from neurophysiological studies in monkeys. Neurons located in several structures (putamen, parietal, premotor areas) of the macaque brain (Rizzolatti et al., 1981; Fogassi et al., 1996; Graziano et al., 1997, 1999; Duhamel et al., 1998) have been shown to respond both to touches delivered on a specific body part (for example the hand or the face) and to visual or auditory stimuli presented close to the same body part. The visual or auditory RF of these neurons is in spatial register with the tactile RF: the neuronal response is greater at shorter distance (~ 5 cm) between the hand and the visual or auditory source, and becomes null when the stimulus is presented far from the body part, that is about 30 cm away. Single-cell studies in monkey have also showed that peripersonal space representation is not fixed, but is plastic changing with experience. In particular, Iriki and colleagues (Iriki et al., 1996; Ishibashi et al., 2000) documented that after the animal had repeatedly used a tool to retrieve distant food, the visual RF of intraparietal visual–tactile neurons was elongated to include the entire length of the tool, whereas originally it was limited to the space around the hand (that is, the visual peri-hand space expanded).

In humans, evidence for the existence of a multisensory system devoted to peripersonal space representation mainly come from neuropsychological studies on cross-modal extinction in right brain damaged (RBD) patients. In such studies (di Pellegrino et al., 1997; Làdavas et al., 1998; Farné and Làdavas, 2002), perception of a tactile stimulus on a contralesional body part (hand or head) was extinguished by a simultaneous visual or auditory stimuli presented *near* (~5 cm) the ipsilesional body part, but not by a visual or auditory stimuli presented far away (~35 cm distance). This pattern of results is in agreement with an integrated multisensory system coding the near space. Due to this system, the visual stimulus presented *near* the ipsilesional body part would activate the somatosensory representation of the corresponding body part, thus extinguishing the contralesional tactile stimulation. Studies on extinction patients also reported behavioral evidence of visual peripersonal space extension due to tool-use. Left tactile extinction normally produced by visual stimuli applied near the right hand, was induced also by visual stimuli applied far from the right hand, near the tip of a right-hand held tool, after the patients used this tool to retrieve objects presented in the far space (Farnè and Làdavas, 2000; Maravita et al., 2001).

Besides studies on extinction patients, other studies in healthy subjects further support the existence of a multisensory peripersonal space in humans, with plastic properties depending on experience. In particular, Holmes et al. (2004, 2007a), by using the cross-modal congruency task, showed a modification of the visual–tactile integrative area of the hand in healthy humans after they actively used a tool.

Two major inferences can be drawn from previous experimental results: (i) Coding of peripersonal space is multisensory, its representation being activated by tactile stimuli as well as by visual or auditory stimuli near the body. Such integrated processing may have a strong value in aiding detection of a stimulus approaching the body, before the contact with the skin occurs, and in preparing an adequate motor response to it. (ii) The coding of space as near (that is as peripersonal), implicating interaction between tactile events with visual (or auditory) events, is not determined only by the distance from the body, but depends also on the relation between the body and the external objects. The use of a tool to extend our effectors, that makes distant objects reachable, seems to promote an extension of peripersonal space, with a remapping of far space as near space.

In the last decades, the problem of space representation has been successfully faced via the computational approach based on artificial neural networks. In particular, in their influential papers, Pouget and colleagues (Pouget and Sejnowski, 1995, 1997; Pouget et al., 2002; Avillac et al., 2005) proposed computational models where neurons in the parietal cortex perform sensorimotor transformation for space representation and multisensory integration, by computing basis functions of their sensory and postural inputs. The basis function approach was also used to simulate some aspects of unilateral spatial neglect in vision modality (Pouget and Sejnowski, 2001). These models have helped to clarify properties of parietal neurons and their role in codifying spatial information. However, these models neglect important issues of spatial representation, such as the segregation between near and far space representation, the attentional competition between the representations of the two hemispaces (as emerge in extinction patients), the plasticity of space representations.

In order to investigate these latter aspects, we recently developed a neural network model of visual–tactile representation of the peripersonal space around the left hand and around the right hand (Magosso et al., 2010a,b). Here, the network has been extended to include auditory modality too. Indeed, although auditory peripersonal space (where auditory and tactile information are integrated) has been principally documented around the head (Graziano et al., 1999; Farné and Làdavas, 2002), in a recent study (Serino et al., 2007) Serino et al. (2007) have shown that an auditory peripersonal space also exists around the hand. In the same work, the authors documented that the auditory peri-hand space exhibits plastic properties –following tool-use – similar to those previously found for the visual peri-hand space. Furthermore, a subsequent study (Bassolino et al., 2010) demonstrated that a visual–tactile integration task performed by the hand also affects the audio–tactile integrative peri-hand space, suggesting that visual and auditory peripersonal space representations share the same integrative multisensory system.

These results motivate the inclusion of the auditory modality in our model. The model proposed here is able to simulate, and explain in terms of neural responses, most of the *in vivo* results delineated above.

## MODEL DESCRIPTION

In this section we describe the structure of the neural network. The network is devoted to mimic the multisensory representation of the peri-hand space – both as to the left hand and right hand – in basal conditions (that is before tool-use), and to simulate tool-use training experiments involving expansion of the peri-hand integrative area. Peripersonal space representation and its plasticity have been simulated both as to a healthy subject and a RBD patient with left tactile extinction. All model equations can be easily derived by referring to our previous works (Magosso et al., 2010a).

### *Structure of the neural network*

The network consists of two subnetworks, reciprocally interconnected, each subnetwork referring to the contralateral hand of a hypothetical subject (**Figure 7**).

The single subnetwork embodies four areas of neurons. The three upstream areas are bidimensional lattices of unimodal neurons, responding, respectively, to tactile stimuli on the contralateral hand (*tactile area*), to visual stimuli (*visual area*) and to auditory stimuli (*auditory area*) on the same hand and around it. Each neuron has its own RF (described via a Gaussian function), through which it receives external stimulation. In all areas, the RFs are in hand-centered coordinates and topologically organized, so that proximal neurons within each area respond to stimuli coming from proximal positions of the hand and space. According to data in the literature (Mickey and Middlebrooks, 2003), we assumed that the RF of auditory neurons is larger than that of the tactile and visual neurons. The tactile area maps a surface of 10 cm × 20 cm, roughly representing the surface of the hand. Both the visual and auditory areas cover a space of 15 cm × 100 cm, representing the space *on* the hand and *around* it (extending by 2.5 cm on each side and 80 cm ahead). Moreover, neurons within each unimodal area interact via *lateral synapses* with a "Mexican hat" arrangement (that is, with short-range excitation and long-range inhibition).

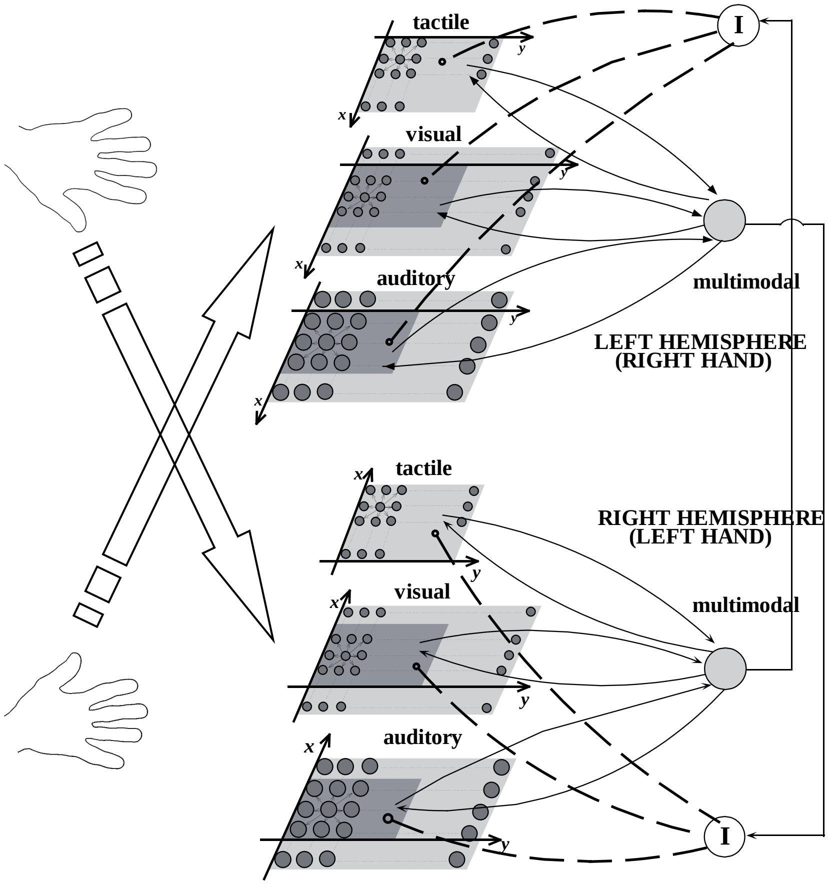

**FIGURE 7 | Layout of the neural network for peri-hand space representation.** The model includes two subnetworks, one per hemisphere, each corresponding to the contralateral hand and surrounding space. Each subnetwork includes three unimodal areas (tactile, visual, and auditory) connected with a downstream multimodal area. The two subnetworks interact via inhibitory interneurons. The gray circles represent excitatory neurons; the continuous arrows linking neurons or areas of neurons denote excitatory connections, the dashed lines denote inhibitory connections. I indicate inhibitory interneurons. Neurons in the auditory areas are made bigger to denote their larger RF with respect to tactile and visual neurons.

The unimodal neurons send *feedforward synapses* to a fourth downstream multimodal area devoted to multisensory representation of peri-hand space. For the sake of simplicity, we considered a single multimodal neuron, covering the entire peri-hand space. Data in the literature, indeed, stresses the existence of multimodal neurons with RF as large as the whole hand (Rizzolatti et al., 1981; Graziano et al., 1997; see also Discussion for such simplification). The tactile feedforward synapses have a uniform distribution. The strength of the visual and auditory feedforward synapses is constant on the hand and decreases exponentially as the distance between the neuron's RF and the hand increases. **Figure 8A** shows the pattern of the feedforward synapses from the three unimodal areas. According to such synapses arrangement, the multimodal neuron has a tactile RF covering the entire hand, and a visual and an auditory RF matching the tactile RF and extending some centimeters around it. **Figure 8B** displays the response of the multimodal neuron in one hemisphere to a visual or auditory stimulus located at different distances from the contralateral hand. The visual or auditory response of the multimodal neuron decreases as the distance between the stimulus and the hand increases, in agreement with neurophysiological data (Graziano et al., 1997, 1999).

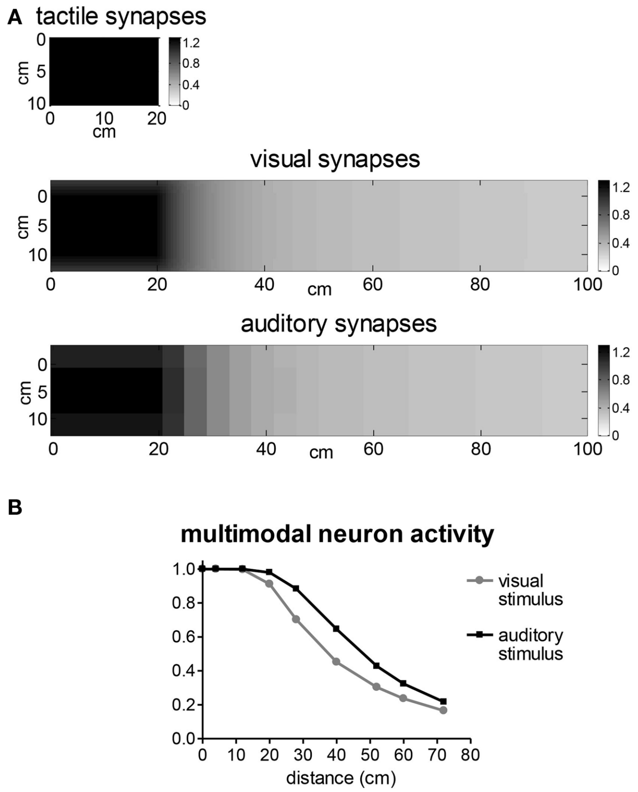

**FIGURE 8 | (A)** Pattern of the feedforward synapses from the tactile, visual and auditory area to the downstream multimodal area in the left hemisphere (for basal parameter values, i.e., healthy subject). The *x* (vertical) and *y* (horizontal) axes represent the coordinates of the RF center of the unimodal neurons; the gray scale indicates the strength of the synapse connection. **(B)** Response of the multimodal neuron in one hemisphere, to a visual or auditory stimulus located at different distances from the corresponding hand (for basal parameter values, i.e., healthy subject). Zero distance means that the stimulus is placed on the hand. Neuron response is normalized with respect to its maximum saturation activity (that is, value one corresponds to the maximal neuron activation).

The multimodal neuron within one hemisphere sends *feedback excitatory synapses* to the upstream unimodal areas in the same hemisphere. The feedback synapses have the same arrangement as the feedforward synapses.

The two hemispheres interact via a competitive mechanism realized by means of inhibitory interneurons. This competition is essential to reproduce data in extinction patients. The inhibitory interneuron in one hemisphere receives information from the multimodal neuron in the other hemisphere and sends inhibitory synapses locally to the unimodal areas. The inhibitory synapses have the same spatial arrangement as the feedback and feedforward synapses.

The input–output relationship of each neuron (unimodal, multimodal, and inhibitory) includes a first-order dynamics and a static sigmoidal relationship. Each neuron is normally in a silent state and can be activated if stimulated by a sufficiently high excitatory input.

### *Parameters of the neural network (healthy subject and RBD patient)*

Basal parameter values were assigned on the basis of neurophysiological and behavioral literature, in order to reproduce a healthy subject. In particular, the healthy subject has been mimicked assuming the same parameter values in the two hemispheres. The RBD patient with left tactile extinction has been reproduced by decreasing the strength of all excitatory synapses (both lateral and feedforward) originating from the tactile unimodal neurons in the right hemisphere (Magosso et al., 2010a,b). This reduction in synaptic strength could reproduce the effect of a reduction – due to the lesion – in the number of effective excitatory neurons which contribute to the activity in that region.

## RESULTS

First, we performed simulations, both in the healthy subject and in the RBD patient, to assess peri-hand space representation in basal conditions (that is before tool-use). To this aim, the network has been stimulated with unilateral or bilateral cross-modal inputs. The incoming stimulus of any modality mimics a quite punctual stimulus. Then, tool-use training has been simulated (by training network synapses, see below) and the extension of the integrative peri-hand area re-evaluated after training (that is after tool-use). Figures show network response at approximately steady-state conditions after stimuli application.

### *Peri-hand space representation before tool-use*

We evaluated whether in the model tactile stimuli on one hand are integrated with stimuli of different modalities (visual or auditory) presented in the space around the same hand, and whether this integration exhibits a near–far modulation, as observed *in vivo* (Serino et al., 2007). To this aim, we applied a weak tactile stimulus on the right hand in isolation (unimodal stimulation) or associated with a concurrent auditory (or visual) stimulus in the same hemispace located near or far from the hand (cross-modal unilateral stimulation). Results are presented in **Figure 9A–C** as to an audio–tactile stimulation. In each plot, the panels show the activity in the tactile area, in the auditory area and in the multimodal area of the stimulated

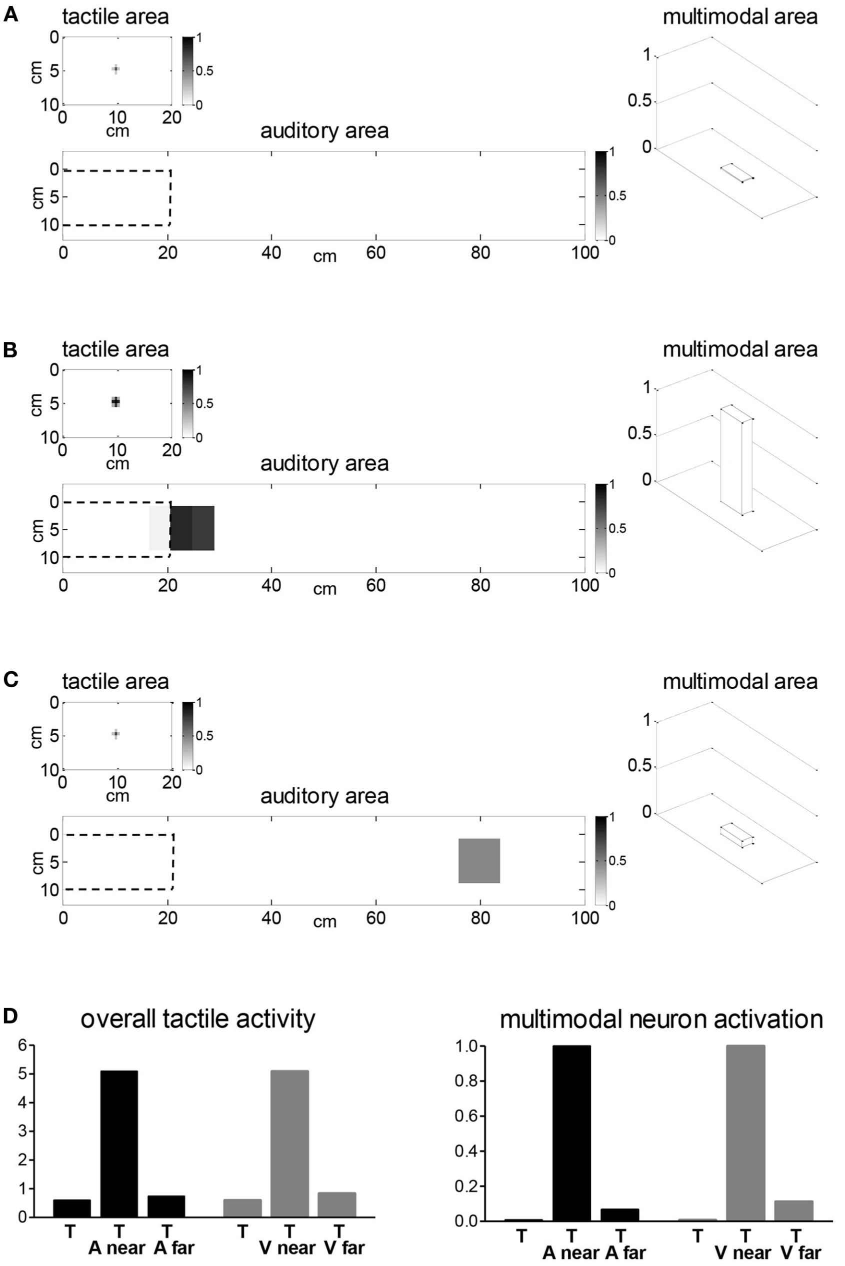

**FIGURE 9 | (A)** Network response to a weak tactile stimulus on the right hand. Plots show activity in tactile and in the auditory area (represented as gray plot) and in the multimodal area (represented via a 3D bar) of the left hemisphere. The dashed border within the auditory area delimits the auditory space on the hand. **(B)** Network response to unilateral cross-modal stimulation with a weak tactile stimulus on the right hand [as in **(A)**] and an auditory stimulus near the same hand. The auditory stimulus is centered at horizontal position $y = 24$ cm (that is, at 4 cm distance from the hand). **(C)** Network response to unilateral cross-modal stimulation with a weak tactile stimulus on the right hand [as in **(A)**] and an auditory stimulus far from the same hand. The auditory stimulus is centered at $y = 80$ cm (that is, at 60 cm distance from the hand). **(D)** Histograms representing overall activation in the tactile area (computed by summing activities of all neurons in that area) and multimodal neuron activation in the left hemisphere in case of the previous audio–tactile stimulations, and in case of visuo-tactile stimulations obtained by replacing the auditory stimulus with a visual stimulus. T, tactile alone; T and A near (or V near), tactile stimulus plus auditory (or visual) near stimulus; T and A far (or V far), tactile stimulus plus auditory (or visual) far stimulus.

hemisphere (the visual area is not shown since it remains in a silent state). In **Figure 9A**, the weak tactile stimulus is presented in isolation: the stimulus produces only a slight activity in the tactile area, unable to activate the corresponding multimodal neuron. In **Figure 9B**, the same tactile stimulus is applied in combination with an auditory stimulus *near* the hand (4 cm apart): in this condition, the activity in the tactile area is significantly enhanced with respect to the previous case. Indeed, the near auditory stimulus activates the multimodal neuron, which in turn, via the feedback synapses, reinforces the tactile activation. It is worth noticing that the auditory input produces a larger activation in the unimodal area with respect to tactile input (as well as visual input, see subsequent results) due to the larger RF of auditory neurons. In **Figure 9C**, the tactile stimulus is combined with a far auditory stimulus (60 cm from the hand). In this case, the far sound produces only a very mild activation of multimodal neuron, because of the weak feedforward synapses (see **Figure 8**), and tactile activation remains unchanged with respect to the unimodal tactile stimulation. Similar results can be obtained by replacing auditory stimuli with visual stimuli, as shown by the histograms in **Figure 9D**. The histograms display the overall activity in the unimodal tactile area and the activation of the multimodal neuron, in the three examined conditions, when using an auditory stimulus or a visual stimulus. According to previous findings, in the model audio–tactile or visuo-tactile integration occurs in the space proximal to the hand, and not in the far space, in agreement with *in vivo* data (Macaluso et al., 2000; Serino et al., 2007).

Then, we investigated how a tactile stimulus on one hand (e.g., the left hand) interacts with a concurrent visual or auditory stimulus in the opposite hemispace (bilateral cross-modal stimulation), and how this interaction may depend on the position of the visual or auditory stimulus with respect to the other hand. We applied such stimulations both in the simulated healthy subject and in the simulated RBD patient.

**Figure 10** displays model results in case of bilateral visuo-tactile and audio–tactile stimulations in the healthy subject, with the tactile input applied on the left hand (right hemisphere), and the visual or auditory input applied in the right hemispace (left hemisphere). **Figure 10A** shows network behavior in case of visuo-tactile stimulation whit the visual input applied *near* the right hand (left hemisphere). Each stimulus produces a cluster of nearby excited neurons (activation bubble) in the corresponding unimodal area, able to trigger, via the feedforward synapses, the related multimodal neuron. The concurrent activation of the two multimodal neurons leads to a competition between the two hemispheres, via the inhibitory interneurons. In this case (healthy subject), the left tactile stimulus and the near right visual stimulus exert a similar excitatory action on the corresponding multimodal neuron (see feedforward synapses in **Figure 8**), and the competition between the two hemispheres is balanced. The final outcome is the coexistence of activations in both hemispheres, with the multimodal neuron maximally activated in each hemisphere.

In **Figure 10B**, the tactile input on the left hand is applied in combination with a visual input *far* from the right hand (60 cm from the hand). The far visual stimulus, because of the weak feedforward synapses (see **Figure 8**), produces only a very mild activation of the corresponding multimodal neuron, whereas right hemisphere multimodal neuron is maximally triggered by the tactile stimulus.

The histogram in **Figure 10C** synthetically describes network responses to bilateral visuo-tactile stimulations, with the right visual stimulus at several different distances from the right hand, by reporting only activity of the multimodal neurons in the two hemispheres. The tactile stimulus always activates the corresponding multimodal neuron. Conversely, due to the pattern of visual feedforward synapses (**Figure 8**), the activation of the left multimodal neuron decreases as the distance of the visual stimulus from the right hand increases. In particular, only near visual stimuli (applied at a distance not greater then ~20 cm from the hand) are able to trigger the multimodal neuron at its maximum level. The model predicts similar results in case of bilateral audio–tactile stimulations, with the auditory input at different distances from the right hand (**Figure 10D**).

The same bilateral visuo-tactile and audio–tactile stimulations as in **Figure 10** have been replicated in the RBD patient simulated by reducing the strength of excitatory synapses emerging from the right hemisphere tactile area (see "Model Description"). Results are reported in **Figure 11**. **Figure 11A** (visuo-tactile stimulation) shows that the *near* right-hand visual stimulus activates the multimodal neuron in the left hemisphere, competing with the simultaneous left tactile stimulus. In this case, since right hemisphere tactile activation is impaired by the lesion, the competition is unbalanced, with the right visual stimulus having a higher competitive strength than the left tactile stimulus. The final outcome is a strong reduction of the activity in the right hemisphere tactile area and a consequent deactivation of the corresponding multimodal neuron. This network response may be interpreted as extinction of left tactile stimulus (see also Discussion). On the contrary, a far visual stimulus (60 cm distance from the right hand, **Figure 11B**) exerts a very weak competition with the left tactile stimulus. As a consequence, tactile activation may emerge despite the deficit, triggering the corresponding multimodal neuron, which in turn reinforces unimodal tactile activity via the feedback synapses. It is worth noticing, indeed, the visible stronger activation in the right hemisphere tactile area with respect to **Figure 11A**. Network response in **Figure 11B** may correspond to perception of the tactile stimulus (see also Discussion). Bilateral visuo-tactile stimulations with the right visual input located at several different positions (histogram in **Figure 11C**) show that deactivation of right hemisphere multimodal neuron (i.e., left tactile extinction) occurs in case of visual stimuli within 30 cm from the hand, and not for more distant stimuli, in agreement with *in vivo* studies of visuo-tactile extinction (Làdavas et al., 1998). Analogous results are predicted by the model by replacing the visual stimulus with the auditory stimulus (**Figure 11D**), in agreement with *in vivo* studies of audio-tactile extinction (Farné and Làdavas, 2002).

#### *Network training (tool-use training)*

A training experiment has been simulated in which the hypothetical subject utilizes a tool with the right hand to interact with visual stimuli (objects) in the far space. The use of the tool by the right hand has been mimicked by applying both a tactile and a visual input to the left hemisphere (see **Figure 12A**). The tactile input represents the portion of the hand stimulated while holding the tool. The visual input represents the region of the visual space functionally relevant for the tool-use, selected, for instance, by top–down attentive mechanisms. Here, we adopted an elongated visual input, that could mimic the use of a rake to retrieve objects from the far space (Iriki et al., 1996; Farnè and Làdavas, 2000), requiring allocation of attention toward a wide portion of the visual space. The auditory input has been set to zero, assuming that in the simulated conditions the auditory information play a minor role during training.

The application of the previous inputs to the network produces the activation of the corresponding regions in the unimodal areas within the left hemisphere, and the activation of the left multimodal neuron. During the application of these inputs, the

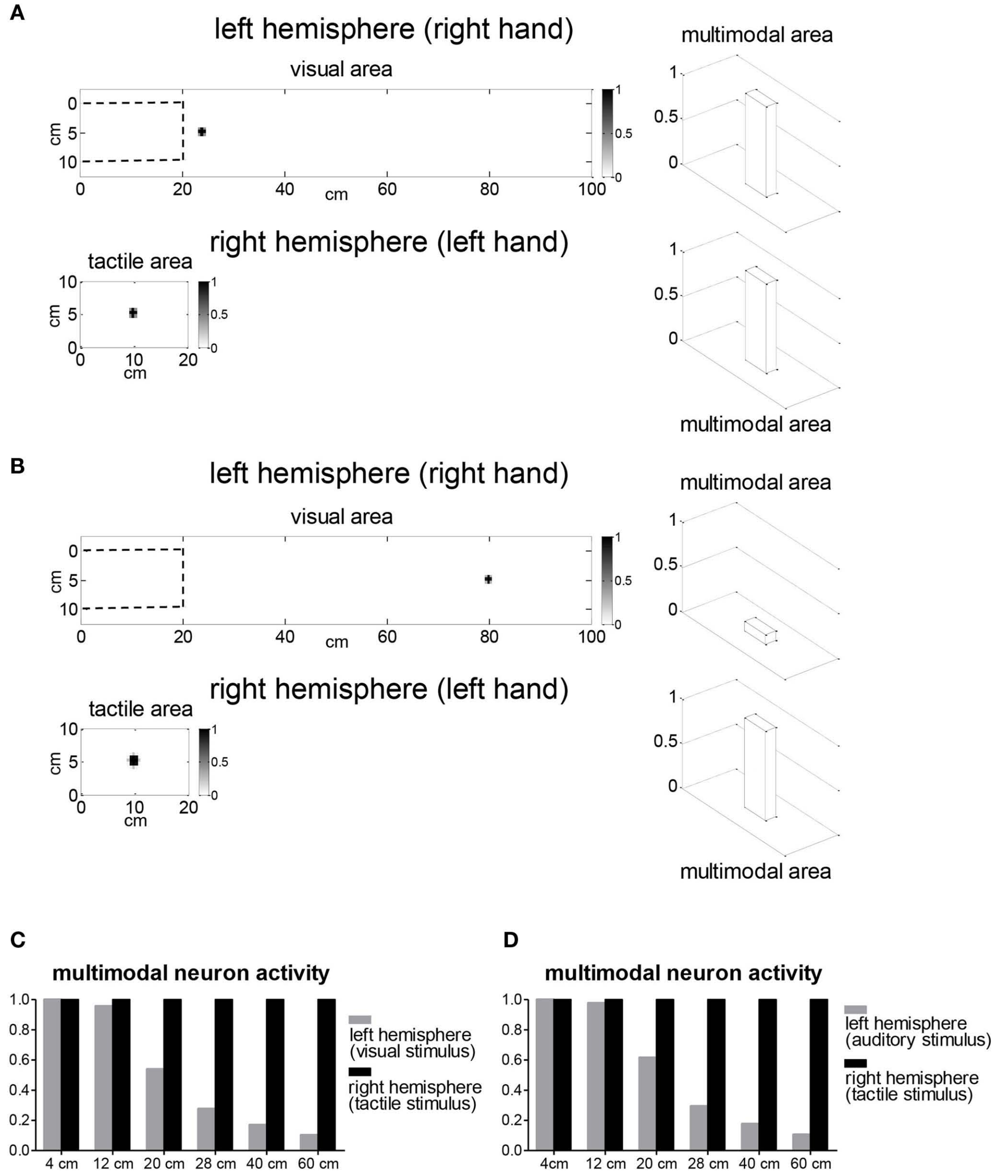

**FIGURE 10 | Network response to bilateral visuo-tactile and audio–tactile stimulations in the healthy subject, with a tactile stimulus on the left hand and a visual or auditory stimulus in the right hemispace. (A)** Activity in the two hemispheres in case of visuo-tactile stimulation with a near right visual stimulus. Plots show the activity in the stimulated unimodal areas and the activity in the multimodal areas, in both hemispheres. The non-stimulated unimodal areas are silent and are not displayed. The dashed border within the visual area delimits the visual space on the hand. The visual stimulus is applied at $y = 24$ cm, that is at 4 cm distance from the right hand. Note that the tactile stimulus is stronger with respect to that applied in **Figure 9**, being able to produce sufficient activation in the tactile area and trigger the downstream multimodal neuron. In these conditions, multimodal neurons in both hemispheres are maximally activated. **(B)** Activity in the two hemispheres in case of a far right visual stimulus, applied at $y = 80$ cm (that is at 60 cm distance from the right hand). In this case, left hemisphere multimodal neuron exhibits only a scarce activation. **(C)** Histogram showing the activation of the multimodal neurons in the two hemispheres in response to visuo-tactile bilateral stimulations, with the visual stimulus located at different distances from the right hand. The first and last positions correspond to the same simulations as **(A)** and **(B)**. **(D)** Histogram showing the activation of the multimodal neurons in the two hemispheres in response to audio–tactile bilateral stimulations, with the auditory stimulus located at different distances from the right hand.

feedforward synapses from unimodal neurons to the multimodal neuron in the left hemisphere have been assumed to modify according to a Hebbian learning rule with an upper saturation: synapses are reinforced in presence of the simultaneous activation of the pre-synaptic and post-synaptic neurons, until a maximal value is reached. Moreover, we hypothesized that synapses *on* the hand are already at their maximum value even before tool-use, since they are frequently and repeatedly involved in the daily perception of the peri-hand space.

The pattern of the visual feedforward synapses after the Hebbian learning are shown in **Figure 12B**: visual synapses reinforce significantly along the extended visual input highlighted during the

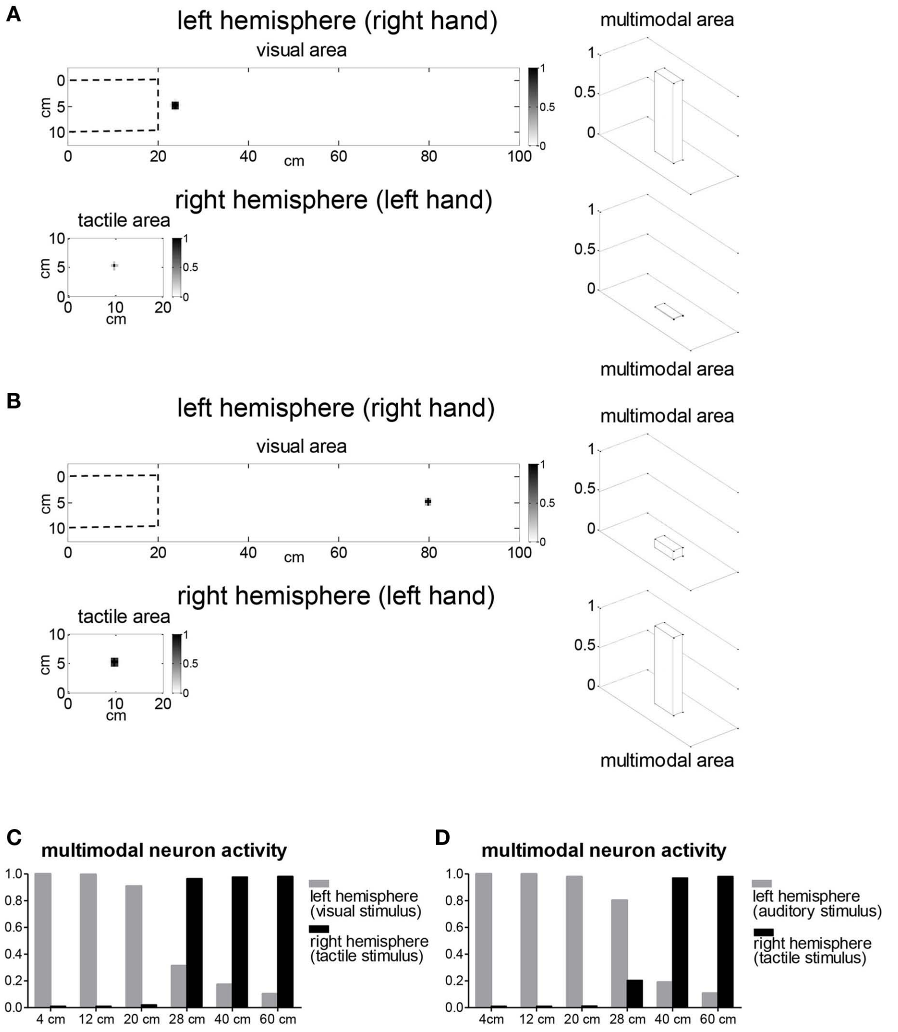

**FIGURE 11 | Network response to bilateral visuo-tactile and audio–tactile stimulations in the RBD patient, with a tactile stimulus on the left hand and a visual or auditory stimulus in the right hemispace.** Stimuli intensity is the same as in **Figure 10**. **(A)** Activity in the two hemispheres in case of a visuo-tactile stimulation with a near right visual stimulus (4 cm distance from the right hand). Left hemisphere multimodal neuron is maximally activated, whereas right hemisphere multimodal neuron is deactivated (left tactile extinction). **(B)** Activity in the two hemispheres in case of a far right visual stimulus (60 cm distance from the right hand). The right visual stimulus produces only a weak activation of the multimodal neuron, and left tactile stimulus is able to maximally trigger the corresponding multimodal neuron. **(C,D)** Histograms showing the activation of the multimodal neurons in the two hemispheres in response to visuo-tactile and audio–tactile bilateral stimulations, with the visual stimulus or auditory stimulus located at different distances from the right hand. Only far right stimuli allow activation of the multimodal neuron by the left touch.

training. Tactile synapses do not modify because of the previous assumptions; auditory synapses do not change since no auditory pre-synaptic activity is present during training.

All equations and parameters concerning model training and plasticity can be found in our previous paper (Magosso et al., 2010a).

### *Peri-hand space representation after tool-use training*

During training, only visual feedforward synapses modify (see previous section). Hence, after network training, just visuo-tactile stimulations (both unilateral and bilateral) have been repeated to evaluate possible modifications of the integrative visuo-tactile peri-hand area. Audio–tactile stimulations have not replicated since they produce the same results as before training.

**Figure 13** shows network response to unilateral visuo-tactile stimulation on the right hand, involving a weak tactile stimulus on the right hand associated with a right visual stimulus near or far from the same hand. At variance with basal conditions (see **Figure 9D**), the far visual stimulus is now able to activate the multimodal neuron and can reinforce tactile activation via to the back projections from the multimodal neuron to the tactile unimodal neurons. That is, the far visual stimulus behaves as the near one.

**Figure 14** reports the results of bilateral visuo-tactile stimulations, with the visual stimulus located at different distances from the right hand, in the healthy subject (**Figure 14A**) and in RBD patient (**Figure 14B**), after tool-use training. In the trained condition, the right visual stimulus located in any of the examined positions activates

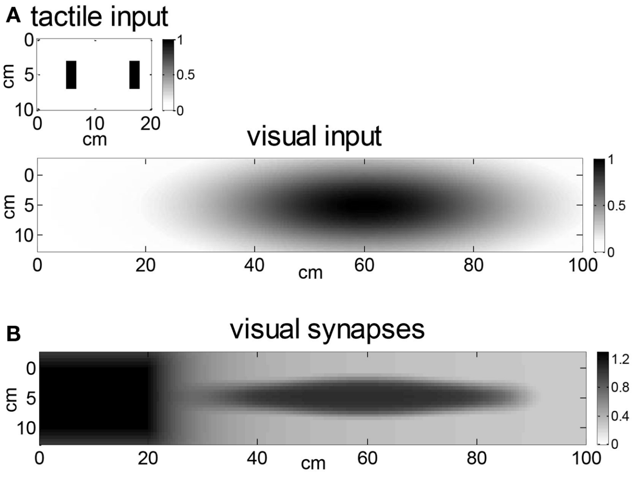

**FIGURE 12 | (A)** Tactile and visual inputs used to simulate tool-use training with the model. These inputs were applied to the left hemisphere since we simulated the use of the tool with the right hand. **(B)** Feedforward synapses from the visual area to the multimodal neuron in the left hemisphere after training, computed via the application of a Hebbian rule during stimulation of the network by the tool-related inputs (compare with **Figure 8A**). Tactile and auditory synapses remain unchanged with respect to before-tool condition (that is, the same as in **Figure 8A**).

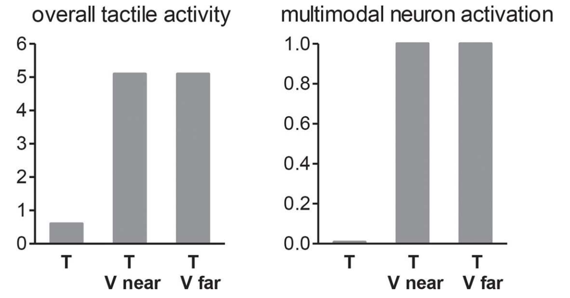

**FIGURE 13 | Histograms represent overall activation in the tactile area and multimodal neuron activation in the left hemisphere after tool-use training in the following cases: (i) single weak tactile stimulation on the right hand, (ii) weak tactile stimulation on the right hand plus visual stimulation near the right hand; (iii) weak tactile stimulation on the right hand plus visual stimulation far from the right hand (i.e., the same stimulations as in Figure 9D)**. The far visual stimulus now behaves as the near one (compare with result in **Figure 9D**).

the multimodal neuron in the left hemisphere (that is, the far space is recoded as near space). In the healthy case, activation of the multimodal neuron triggered by the right visual stimulus coexists with activation of the multimodal neuron boosted by the left tactile stimulus. In the patient, inhibition of the left tactile stimulus (i.e., deactivation of the right hemisphere multimodal neuron) occurs not only for near right visual stimuli but also for visual stimuli in the more distant space.

## DISCUSSION

We implemented a neural network with limited complexity, including three unimodal areas (visual, tactile, auditory) and a multimodal area connected via excitatory feedforward and feedback synapses within each hemisphere, and a competitive interaction via inhibitory interneurons between the two hemispheres.

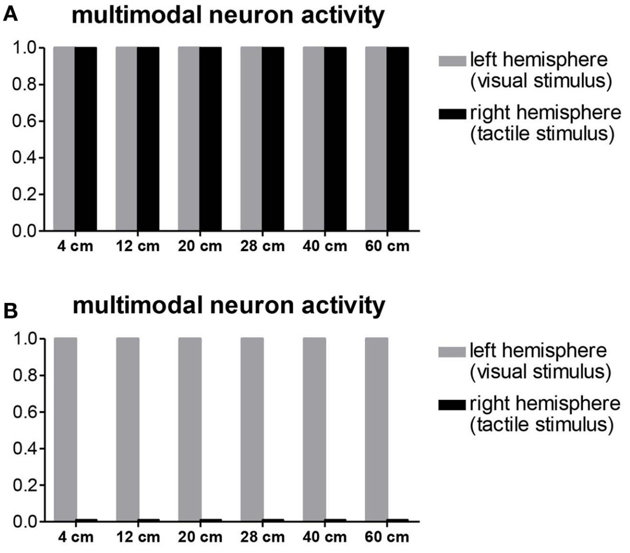

**FIGURE 14 | Activity of the multimodal neurons in the two hemispheres in response to bilateral visuo-tactile stimulations with tactile stimulus on the left hand and visual stimulus at different distances from the right hand (as in Figures 10 and 11) applied to the healthy subject (A) and to the RBD patient (B) after tool-use training with the right hand.** The visual stimulus in any position activates the corresponding multimodal neuron, which, in case of the healthy subject **(A)** coexists with the simultaneous activation of the right hemisphere multimodal neuron, whereas, in case of the RBD patient **(B)**, inhibits the activation of the right hemisphere multimodal neuron.

Model architecture has several physiological counterparts. (i) The multimodal neurons may correspond to cells in the parietal and frontal cortex (observed via electrophysiological measures in animals; Fogassi et al., 1996; Graziano et al., 1997, 1999, and via neuroimaging studies in humans; Bremmer et al., 2001; Makin et al., 2007) having visual, auditory, and tactile RFs in spatial register and matching specific body parts. (ii) The upstream unimodal layers may account for primary and secondary unisensory areas, which project into the multisensory areas through different pathways (Graziano et al., 1997, 1999; Duhamel et al., 1998). (iii) The presence of back projections from the multimodal neuron into the upstream unimodal areas is supported by recent data according to which response in a unimodal area may be modulated by stimulation in a second modality (Driver and Spence, 2000; Macaluso et al., 2000; Macaluso and Driver, 2005). (iv) The existence of an interhemispheric competition for accessing limited attentional resources in peripersonal space has received striking evidence from studies on extinction patients (Hillis et al., 2006).

Some simplifications included in the model deserve a few comments. A first important simplification is the use of a single unit at the multisensory level. This unit represents a pool of neurons having similar RF that covers the entire peri-hand space. Activation of this neuron signals the involvement of the peripersonal space regardless of the specific spatial location of the stimulus within that space (that is the multisensory unit is spatially unspecific within the peripersonal space). This simplification is justified since here we aim at reproducing facilitatory and inhibitory cross-modal interactions (mediated by the multisensory layer) that do not depend strictly on the specific spatial locations of the stimuli, provided the stimuli are applied within the peri-hand spaces. This also justifies the pattern of the inhibitory mechanism implemented in the model: inhibition affects unisensory neurons coding the space on or near the hand, i.e., is tuned for the peripersonal space, but is not spatially selective within the peripersonal space. Of course, other cross-modal phenomena more related with spatial localization and resolution (such as the visual enhancement of touch and the ventriloquism effect), could involve different mechanisms, e.g., direct interactions (excitatory and inhibitory) among the unisensory areas. However, these phenomena are behind the aim of the proposed model, hence we avoided to include other mechanisms here.

Another important oversimplification of the present model is that the spatial arrangement of visual, auditory, and tactile RFs of the unimodal neurons has been set *a priori* on and around the hand; that is we avoid considering explicitly the problem of coordinate transformations between different reference frames (e.g., from eye-centered to hand-centered coordinates), a problem widely investigated in other studies by means of neural network models (Pouget et al., 2002; Avillac et al., 2005). We claim that such simplification mainly reduces model complexity, without affecting model results and inferences.

The model is able to reproduce a variety of results concerning peripersonal space representation and its plastic modifications; in the following, we will highlight how the model may help interpretation of *in vivo* data, rises new questions and inspires novel experiments on the basis of the generated predictions, and potentially promotes advancement in the clinical practice involving multisensory integration.

### *Multimodal neurons and behavioral responses in the healthy subject and RBD patient*

A first important point is that the model is able to relate behavioral results with neural responses. In the model, activation of the multimodal neuron signals the involvement of the peri-hand space triggered by a tactile stimulus on the hand, or by a visual or auditory stimulus near the hand. The two multimodal neurons compete via inhibitory mechanisms in responding to stimuli in the contralateral sides of peripersonal space (Hillis et al., 2006). The final outcome of this competition may be coexistence of activation of both multimodal neurons or prevalence of one hemisphere over the other.

In case of bilateral cross-modal (visuo-tactile or audio–tactile) stimulation applied to the healthy subject, the model predicts the coexistence of both multimodal neurons activations, when the visual or auditory stimulus is applied in the near space (see **Figure 10**). This model result mimics the balanced allocation of resources toward the two peripersonal hemispaces in the healthy subject, and reproduces the capability of a healthy subject to perceive and report right-hand and left-hand stimulations applied simultaneously (Hillis et al., 2006).

Furthermore, the model has been used to interpret extinction of left tactile stimuli in bilateral stimulation trials (that is extinction *across* hemispaces) observed in right brain damage patients. The patient has been simulated by reducing the strength of the excitatory synapses emerging from the right-hemisphere tactile area. In our model, tactile unisensory area does not correspond to primary somatosensory cortex; rather it reflects different stages of somatosensory processing, involving also higher-level somatosensory areas (such as second somatosensory area), that may be compromised by the lesion (Sarri et al., 2006). Hence, the modification assumed in the model aims at accounting for the damage (e.g., loss of neurons) of higher-level somatosensory areas in the parietal cortex. It is worth noticing that with this alteration, the network is able to replicate the preserved ability to detect isolate contralesional stimuli in the RBD patient. Indeed, in absence of a simultaneous competition with the right hand representation (e.g., in case of isolated left tactile stimuli, or in case of a simultaneous far visual or auditory stimulation, see **Figure 11**), the left tactile stimulus – despite the damage – is able to trigger the corresponding multimodal neuron. This may correspond to conscious perception of the tactile stimulus, and reproduces the preservation of tactile sensation in the patient. Conversely when a competition with right peripersonal space occurs (because of a simultaneous right tactile stimulus, or visual or auditory stimulus near the right hand, **Figure 11**), a weak activity still survives in the right tactile area, but it is insufficient to excite the multimodal neuron which is completely deactivated. This result can correspond to left tactile extinction, that is unawareness of left tactile stimulus. These model outcomes are supported by recent ERP and fMRI data in tactile extinction patients showing that missed left touches can still lead to an activation of the right somatosensory cortex, but fail to activate the right parietal and frontal cortices (corresponding to the downstream multimodal area in the model), which conversely are activated by consciously perceived left touches (Eimer et al., 2002; Sarri et al., 2006).

### *Implications on extinction patients*

We assumed an impairment in the tactile area of the damage hemisphere (right hemisphere in our simulations) that biases the competition in favor of the healthy (left) hemisphere (that is, in favor of the ipsilesional stimulus). In this way, the model can reproduce unimodal (tactile–tactile) and cross-modal (visuo-tactile or auditory–tactile) extinction across hemispaces (in this paper we do not show results of tactile–tactile extinction since are similar to cross-modal extinction of **Figure 11**)

Some papers (Gainotti et al., 1989, 1990; Costantini et al., 2007) reported that in unilateral brain damage patients, extinction may occur not only across hemispaces, but also within the same hemispace (omission of one stimulus in case of double simultaneous stimulation on the same side of space). Extinction within a single hemispace can be cross-modal or unimodal, and it has been observed both on the side contralateral to the lesion and – although to a much lesser extent – on the side ipsilateral to the lesion.

Extinction within a single hemispace may be explained via competitive mechanisms within the same hemisphere. The present model realizes within-modality competition inside each hemisphere, via lateral inhibition among the unimodal neurons. Accordingly, it may reproduce unimodal extinction inside a single hemispace. In basal conditions, inhibition is weak, and the response to one stimulus (let's say a tactile stimulus) is depressed, but not totally suppressed, by a second tactile stimulus applied in a different position of the same side of space. Total suppression can be reproduced simulating impaired conditions, as in patients, that create a bias in favor of one stimulus at expenses of the other. For example, by strongly increasing the lateral inhibition within the unimodal layer, a very small difference in the intensity of the two stimuli (as it occurs in real stimulation), would produce the survival only of the slightly stronger one. The increased lateral inhibition could simulate a general reduction in the ability of attending to external stimulation, a mechanism that has been hypothesized to underlie also extinction within the ipsilesional hemispace (Gainotti et al., 1989, 1990).

Conversely, this model is not able, in its present version, to replicate cross-modal extinction within the same hemispace since inhibitory competition among different modalities occurs only across the hemispheres and not within the same hemisphere. To reproduce that result the model should be modified, by considering a multimodal layer that includes multiple units, each codifying a part of the peri-hand space, and connecting these units via lateral excitatory and inhibitory synapses. Indeed, the model presented in the Section "Audio–Visual Integration in Superior Colliculus: A Neural Network Model" (model of SC) can predict both unimodal extinction within a single hemispace (thanks to lateral inhibition inside unimodal areas) and cross-modal extinction within a single hemispace ascribing it to the presence of inhibitory lateral synapses within the SC (multimodal) layer.

Identification of the potential functional alterations in the neural circuitry able to explain extinction phenomena, is of relevance not only to improve the knowledge of the neural correlates of that pathological sign, but also to suggest new strategies of rehabilitation. In particular, the model predicts (see **Figure 9**) that the inability of a weak tactile activation to trigger the multimodal neuron may be compensated by a spatially coherent visual or auditory stimulus (that is, near the tactilely stimulated body part). The addition of this stimulus activates the multimodal neuron, which – thanks to the back projections – reinforces tactile activation. This multisensory integration capability may be exploited not only for a short-term improvement of tactile perception, but also for a long-term recovery of somatosensation in patients with tactile extinction. Systematic visuo-tactile (or audio–tactile) stimulation of the pathological side in extinction patients might promote a Hebbian reinforcement of the feedforward synapses (from tactile area to multimodal area) in the damaged hemisphere, that could be effective to re-equilibrate – in a long-lasting way – the competition among the two hemispheres.

#### *Neural correlates of peripersonal space plasticity*

The model is able to simulate re-sizing of peripersonal space after tool-use. In the present study, we simulated only visual peripersonal space expansion. Expansion of auditory peripersonal space may be obtained in a similar way by simulating an auditory–tactile training task. The model attributes the expansion of visual peripersonal space to a reinforcement of visual synapses entering into the multimodal area, which extends the visual RF of multimodal neurons. This hypothesis is supported by recent electrophysiological studies on monkeys (Hihara et al., 2006), and has received further validation in our previous work (Magosso et al., 2010a). The model predicts that, after training, a visual stimulus placed in the space highlighted during the training is able to trigger the corresponding multimodal neuron. In particular, after the training, visual stimuli even far from the right trained hand behave as near visual stimuli. Accordingly, a far right visual stimulus is now able to interact with a tactile stimulus on the same hand, enhancing weak tactile activation (see **Figure 13**). A similar result as to audio–tactile interaction was obtained experimentally in healthy subjects, after they were trained to use a tool to explore the far space in dark conditions (Serino et al., 2007). In the RBD patient, the model predicts that extinction of left touch (that is deactivation of right hemisphere multimodal neuron and reduction of unimodal tactile activity) is no longer modulated by the distance of the right visual stimulus, but occurs in case of both near and far visual stimuli (**Figure 14B**), in agreement with psychophysical data (Farnè and Làdavas, 2000; Maravita et al., 2001).

It is worth noticing that the present model has been used mainly to simulate experiments performed in extinction patient, where the visual peripersonal space is assessed before tool-use and then after tool-use via cross-modal bilateral stimulation. Conversely, the model has not been used here to simulate the relevant results on tool-use plasticity obtained on healthy subjects by Holmes et al using the cross-modal congruency task (Holmes et al., 2004, 2007a,b). Simulation of such task would require the inclusion of several additional aspects (such as representation of target and distractor stimuli, discrimination between target locations on the hand), and for this reason we avoided to consider it here. Model extensions may be performed in subsequent works, in order to replicate also these results.

A number of hypotheses have been included in the model to reproduce the re-sizing of the integrative visuo-tactile area following tool-use. These hypotheses generate some predictions: such predictions can be verified with respect to *in vivo* results, or may suggest novel experiments that can be used for validation or rejection of the underlying hypotheses.

(1) In the model, the change of visual RF of the multimodal neuron critically depends upon the visual input used during the learning phase (see **Figure 12**); the latter may represent the region of the space selected by attentive mechanisms during the training task. Hence, according to the model, different tasks, that require to allocate attention toward different regions of the visual space (e.g., retrieving objects, pressing far buttons with the tip, sorting objects in the far space, etc), should produce different re-sizing of the peri-hand visual–tactile space (for example the formation of a novel integrative peri-hand area at the tip of the tool rather than an elongation along the tool axis). A preliminary validation of this model prediction comes from results of recent studies on extinction patients (Farnè et al., 2005, 2007) and healthy subjects (Holmes et al., 2004). These results suggest a different modification of the boundaries of the visual peripersonal space depending on the region of space where tool-use activity is exerted during training (formation of a novel integrative area at the tip of the tool following a pushing botton task; Holmes et al., 2004; expansion of the peripersonal space along all the length of the tool following a retrieving object task; Farnè et al., 2005, 2007).

(2) During the training, the Hebbian learning rule has been applied only to the feedforward synpases linking active visual neurons to multimodal neuron, whereas feedback synapses from the multimodal neuron toward the active visual neurons have been assumed to remain unchanged. This assumption has two main inferences. The first is that after tool-use, for example with the right hand, a visual stimulus far from the right

hand (in the space highlighted during the training) should facilitate the perception of a weak tactile stimulus on the same hand (as in **Figure 13**). The second is that the reverse should not hold, that is tactile stimuli should not be able to facilitate perception of weak visual stimuli in the far space (because of the weak non-trained feedback synapses from the multimodal neuron to visual neurons coding the far space).

(3) By adopting the classical Hebbian rule (requiring co-occurrence of pre-synaptic and post-synaptic activity), and by applying only visual and tactile inputs during the training (without any auditory inputs), the model predicts an extension of the visual peripersonal space without any modification of the auditory peripersonal space. Of course the reverse would hold in case of replacing visual input with auditory input during training. Experiments could be designed in order to assess whether the training with a specific modality (e.g., visual) extends peripersonal space only in that modality or whether the expansion is transferred to the other modality too (auditory). A preliminary result supporting a shift of peripersonal space expansion from one modality (visual) to another (auditory) is provided by a recent study (Bassolino et al., 2010). To reproduce this shift, some other mechanisms (e.g., direct connections among unisensory areas, whose existence is provided by some recent studies (Macaluso and Driver, 2005; Schroeder and Foxe, 2005)) or modifications of the learning rule should be included in the model.

In conclusions, the present model suggests plausible network topology and neural mechanisms responsible for multisensory representation of peripersonal space; identifies alterations in network nodes and connections able to explain psychophysical results in extinction patients; proposes a biological learning rule able to reproduce the dynamic properties of peripersonal space representation and to provide an explanation of the neural basis of tool-use behavior.

## GENERAL CONCLUSIONS

In conclusion of this paper, we wish to underline some basic ideas and fundamental mechanisms, which emerge from the previous two models.

Although devoted to different problems and simulating different brain regions (the SC in the first model, associative parietal cortex and premotor cortex in the second), the proposed models share some common mechanisms that are briefly summarized below:

(1) *Lateral excitation and inhibition*: Short-range excitation and long-range inhibition among neurons, with a spatial function similar to that of a Mexican hat, is a pattern of connectivity that is ubiquitous in the cortex (Rolls and Treves, 1998). It guarantees: (i) that a single stimulus is represented in a robust manner, being coded by a group of mutually excited units and not by a single cell; (ii) that an incongruent stimulus may be suppressed or eliminated by a proximal stronger stimulus

(2) *Non-linear (sigmoid-like) input–output response*: This kind of response is fundamental to regulate the degree of integration among different stimuli and favor enhancement in the presence of weak individual stimuli.

(3) *Competitive inhibitory mechanisms among different areas*: Competitive mechanisms in processing perceptual inputs may have important functions. They may be essential to select only the most relevant and potentially dangerous stimulus in case of limited resources for attending and responding to external stimuli or to select the neural processing pathway that guarantees a better response to the incoming input.

(4) *Feedback from multisensory to unisensory areas*: Our models assume that the multisensory representation sends a feedback to the upstream unisensory areas (see also Driver and Spence, 2000; Macaluso and Driver, 2005). In view of this feedback, a unisensory representation can be influenced by the other unisensory representations with the occurrence of interesting cross-talk effects. This is fundamental to implement reinforcement of unimodal perception by a cross-modal stimulation when the information provided by one modality is weak (see for example **Figure 9B**) or to resolve ambiguities when information from different modalities are in conflict, merging them into a robust percept (e.g., the ventriloquism phenomenon in case of audio–visual discrepancy).

(5) *Parameter changes*: Parameters in the model can be modified, altering network nodes and connections, to simulate individual variability and/or pathological conditions. The potentialities of this approach are evident for what concerns the study of neuroclinical problems: by simulating the lesioned model, we can provide insight into the neural mechanisms at the basis of psychophysical and behavioral deficits following specific brain lesions.

(6) *Synaptic plasticity*: Certainly, the more distinctive and intriguing feature of an artificial neural network is that – like the actual brain – it can learn from the external environment, shaping its connections on the basis of previous experience, in order to behave in a manner functionally relevant with respect to its environment. The two presented models offer excellent examples of these possibilities, the first demonstrating how multisensory integration capabilities in SC can progressively maturate in a multisensory environment, the second showing how the peripersonal space representation may be plastic and modified by practice.

An interesting aspect, which deserves further studies, is whether these mechanisms (or similar ones) can be effective also in other multisensory structures of the brain, and can be exploited to reach a more general comprehension of how a structure can adapt to a complex multisensory non-stationary external world.

Finally, we wish to stress that this work exemplarily illustrates how theoretical studies based on modeling may complement experimental research to promote advancement in the comprehension of cognitive processes and, specifically, multisensory integration processes. On one hand, empirical results are fundamental to build the mathematical model, identifying model structure, and components. On the other hand, models are fundamental to synthesize the data into a unitary quantitative theory, to explain the specific impact of the involved neural mechanisms on behavior, to generate new predictions and inspire novel related experiments.

## REFERENCES

Alvarado, J. C., Stanford, T. R., Rowland, B. A., Vaughan, J. W., and Stein, B. E. (2009). Multisensory integration in the superior colliculus requires synergy among corticocollicular inputs. *J. Neurosci.* 29, 6580–6592.

Alvarado, J. C., Stanford, T. R., Vaughan, J. W., and Stein, B. E. (2007). Cortex mediates multisensory but not unisensory integration in superior colliculus. *J. Neurosci.* 27, 12775–12786.

Amari, S. (1989). "Dynamical stability of pattern formation of cortical maps," in *Dynamic Interactions in Neural Networks: Models and Data*, eds M. A. Arbib and S. Amari (New York: Springer), 15–34.

Anastasio, T. J., and Patton, P. E. (2003). A two-stage unsupervised learning algorithm reproduces multisensory enhancement in a neural network model of the corticotectal system. *J. Neurosci.* 23, 6713–6727.

Anastasio, T. J., Patton, P. E., and Belkacem-Boussaid, K. (2000). Using Bayes rule to model multisensory enhancement in the superior colliculus. *Neural. Comput.* 12, 1165–1187.

Avillac, M., Denève, S., Olivier, E., Pouget, A., and Duhamel, J. R. (2005). Reference frames for representing visual and tactile locations in parietal cortex. *Nat. Neurosci.* 8, 941–949.

Bassolino, M., Serino, A., Ubaldi, S., and Làdavas, E. (2010). Everyday use of the computer mouse extends peripersonal space representation. *Neuropsychologia* 48, 803–811.

Bremmer, F., Schlack, A., Shah, N. J., Zafiris, O., Kubischik, M., Hoffmann, K., Zilles, K., and Fink, G. R. (2001). Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalencies between humans and monkeys. *Neuron* 29, 287–296.

Calvert, G. A. (2001). Cross-modal processing in the human brain: insights from functional neuroimaging studies. *Cereb. Cortex* 11, 1110–1123.

Colonius, H., and Diederich, A. (2004). Why aren't all deep superior colliculus neurons multisensory? A Bayes' ratio analysis. *Cogn. Affect. Behav. Neurosci.* 4, 344–353.

Costantini, M., Bueti, D., Pazzaglia, M., and Aglioti, S. M. (2007). Temporal dynamics of visuo-tactile extinction within and between hemispaces. *Neuropsychology* 21, 242–250.

Cuppini, C., Ursino, M., Magosso, E., Rowland, B. A., and Stein, B. E. (2010). An emergent model of multisensory integration in superior colliculus neurons. *Front. Integr. Neurosci.* 4:6. doi: 10.3389/fnint.2010.00006

di Pellegrino, G., Làdavas, E., and Farnè, A. (1997). Seeing where your hands are. *Nature* 388, 730.

Driver, J., and Spence, C. (1998). Crossmodal links in spatial attention. *Philos. Trans. R. Soc. Lond. B* 353, 1319–1331.

Driver, J., and Spence, C. (2000). Multisensory perception: beyond modularity and convergence. *Curr. Biol.* 10, R731–R735.

Duhamel, J. R., Colby, C. L., and Goldberg, M. E. (1998). Ventral intraparietal area of the macaque: congruent visual and somatic response properties. *J. Neurophysiol.* 79, 126–136.

Edwards, S. B., Ginsburgh, C. L., Henkel, C. K., and Stein, B. E. (1979). Sources of Subcortical Projections to the Superior Colliculus in the Cat. *J. Comp. Neurol.* 184, 309–330.

Eimer, M., Maravita, A., van, V. J., Husain, M., and Driver, J. (2002). The electrophysiology of tactile extinction: ERP correlates of unconscious somatosensory processing. *Neuropsychologia* 40, 2438–2447.

Eimer, M., and Van Velzen, J. (2002). Crossmodal links in spatial attention are mediated by supramodal control processes: evidence from event-related potentials. *Psychophysiology* 39, 437–449.

Ernst, M. O., and Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends Cogn. Sci.* 8, 162–169.

Farnè, A., Iriki, A., and Làdavas, E. (2005). Shaping multisensory action-space with tools: evidence from patients with cross-modal extinction. *Neuropsychologia* 43, 238–248.

Farnè, A., and Làdavas, E. (2000). Dynamic size-change of hand peripersonal space following tool use. *Neuroreport* 11, 1645–1649.

Farné, A., and Làdavas, E. (2002). Auditory peripersonal space in humans. *J. Cogn. Neurosci.* 14, 1030–1043.

Farnè, A., Serino, A., and Làdavas, E. (2007). Dynamic size-change of peri-hand space following tool-use: determinants and spatial characteristics revealed through cross-modal extinction. *Cortex* 43, 436–443.

Fogassi, L., Gallese, V., Fadiga, L., Luppino, G., Matelli, M., and Rizzolatti, G. (1996). Coding of peripersonal space in inferior premotor cortex (area F4). *J. Neurophysiol.* 76, 141–157.

Frassinetti, F., Bolognini, N., Bottari, D., Bonora, A., and Làdavas, E. (2005). Audiovisual integration in patients with visual deficit. *J. Cogn. Neurosci.* 17, 1442–1452.

Frassinetti, F., Bolognini, N., and Làdavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Exp. Brain Res.* 147, 332–343.

Gainotti, G., De Bonis, C., Daniele, A., and Caltagirone, C. (1989). Contralateral and ipsilateral tactile extinction in patients with right and left focal brain damage. *Int. J. Neurosci.* 45, 81–89.

Gainotti, G., Giustolisi, L., and Nocentini, U. (1990). Contraletaral and ipsilateral disorders of visual attention in patients with unilateral brain damage. *J. Neurol. Neurosurg. Psychiatry* 53, 422–426.

Graziano, M. S., Hu, X. T., and Gross, C. G. (1997). Visuospatial properties of ventral premotor cortex. *J. Neurophysiol.* 77, 2268–2292.

Graziano, M. S., Reiss, L. A., and Gross, C. G. (1999). A neuronal representation of the location of nearby sounds. *Nature* 397, 428–430.

Haggard, P., Christakou, A., and Serino, A. (2007). Viewing the body modulates tactile receptive fields. *Exp. Brain Res.* 180, 187–193.

Hihara, S., Notoya, T., Tanaka, M., Ichinose, S., Ojima, H., Obayashi, S., Fujii, N., and Iriki, A. (2006). Extension of corticocortical afferents into the anterior bank of the intraparietal sulcus by tool-use training in adult monkeys. *Neuropsychologia* 44, 2636–2646.

Hillis, A. E., Chang, S., Heidler-Gary, J., Newhart, M., Kleinman, J. T., Davis, C., Barker, P. B., Aldrich, E., and Ken, L. (2006). Neural correlates of modality-specific spatial extinction. *J. Cogn. Neurosci.* 18, 1889–1898.

Holmes, N. P., Calvert, G. A., and Spence, C. (2004). Extending or projecting peripersonal space with tools? Multisensory interactions highlight only the distal and proximal ends of tools. *Neurosci. Lett.* 372, 62–67.

Holmes, N. P., Calvert, G. A., and Spence, C. (2007a). Tool use changes multisensory interactions in seconds: evidence from the crossmodal congruency task. *Exp. Brain Res.* 183, 465–476.

Holmes, N. P., Sanabria, D., Calvert, G. A., and Spence, C. (2007b). Tool-use: capturing multisensory spatial attention or extending multisensory peripersonal space? *Cortex* 43, 469–489.

Huerta, M. F., and Harting, J. K. (1984). "The mammalian superior colliculus: studies of its morphology and connections," in *Comparative Neurology of the Optic Tectum*, ed. H. Vanega (New York: Plenum), 687–773.

Iriki, A., Tanaka, M., and Iwamura, Y. (1996). Coding of modified body schema during tool use by macaque postcentral neurones. *Neuroreport* 7, 2325–2330.

Ishibashi, H., Hihara, S., and Iriki, A. (2000). Acquisition and development of monkey tool-use: behavioral and kinematic analyses. *Can. J. Physiol. Pharmacol.* 78, 958–966.

Jiang, W., Jiang, H., Rowland, B. A., and Stein, B. E. (2007). Multisensory orientation behavior is disrupted by neonatal cortical ablation. *J. Neurophysiol.* 97, 557–562.

Jiang, W., Jiang, H., and Stein, B. E. (2006). Neonatal cortical ablation disrupts multisensory development in superior colliculus. *J. Neurophysiol.* 95, 1380–1396.

Jiang, W., Wallace, M. T., Jiang, H., Vaughan, J. W., and Stein, B. E. (2001). Two cortical areas mediate multisensory integration in superior colliculus neurons. *J. Neurophysiol.* 85, 506–522.

Kadunce, D. C., Vaughan, J. W., Wallace, M. T., Benedek, G., and Stein, B. E. (1997). Mechanisms of within- and cross-modality suppression in the superior colliculus. *J. Neurophysiol.* 78, 2834–2847.

Kadunce, D. C., Vaughan, J. W., Wallace, M. T., and Stein, B. E. (2001). The influence of visual and auditory receptive field organization on multisensory integration in the superior colliculus. *Exp. Brain. Res.* 139, 303–310.

Làdavas, E., di Pellegrino, G., Farnè, A., and Zeloni, G. (1998). Neuropsychological evidence of an integrated visuotactile representation of peripersonal space in humans. *J. Cogn. Neurosci.* 10, 581–589.

Macaluso, E., and Driver, J. (2005). Multisensory spatial interactions: a window onto functional integration in the human brain. *Trends Neurosci.* 28, 264–271.

Macaluso, E., Frith, C. D., and Driver, J. (2000). Modulation of human visual cortex by crossmodal spatial attention. *Science* 289, 1206–1208.

Magosso, E., Cuppini, C., Serino, A., di Pellegrino, G., and Ursino, M. (2008). A theoretical study of multisensory integration in the superior colliculus by a neural network model. *Neural. Netw.* 21, 817–829.

Magosso, E., Ursino, M., di Pellegrino, G., Làdavas, E., and Serino, A. (2010a). Neural bases of peri-hand space plasticity through tool-use: insights from a combined computational-experimental approach. *Neuropsychologia* 48, 812–830.

Magosso, E., Zavaglia, M., Serino, A., di Pellegrino, G., and Ursino, M. (2010b). Visuotactile representation of peripersonal space: a neural network study. *Neural. Comput.* 22, 190–243.

Makin, T. R., Holmes, N. P., and Zohary, E. (2007). Is that near my hand? Multisensory representation of peripersonal space in human intraparietal sulcus. *J. Neurosci.* 27, 731–740.

Maravita, A., Husain, M., Clarke, K., and Driver, J. (2001). Reaching with a tool extends visual-tactile interactions into far space: evidence from cross-modal extinction. *Neuropsychologia* 39, 580–585.

Martin, J. G., Meredith, M. A., and Ahmad, K. (2009). Modeling multisensory enhancement with self-organizing maps. *Front. Comput. Neurosci.* 3:8. doi: 10.3389/neuro.10.008.2009

Mascaro, M., Battaglia-Mayer, A., Nasi, L., Amit, D. J., and Caminiti, R. (2003). The eye and the hand: neural mechanisms and network models for oculomanual coordination in parietal cortex. *Cereb. Cortex* 13, 1276–1286.

Meredith, M. A., and Stein, B. E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *J. Neurophysiol.* 56, 640–662.

Meredith, M. A., and Stein, B. E. (1996). Spatial determinants of multisensory integration in cat superior colliculus neurons. *J. Neurophysiol.* 75, 1843–1857.

Mickey, B. J., and Middlebrooks, J. C. (2003). Representation of auditory space by cortical neurons in awake cats. *J. Neurosci.* 23, 8649–8663.

Patton, P. E., and Anastasio, T. J. (2003). Modelling cross-modal enhancement and modality-specific suppression in multisensory neurons. *Neural. Comput.* 15, 783–810.

Patton, P. E., Belkacem-Boussaid, K., and Anastasio, T. J. (2002). Multimodality in the superior colliculus: an information theoretic analysis. *Brain Res. Cogn. Brain Res.* 14, 10–19.

Perrault, T. J. Jr., Vaughan, J. W., Stein, B. E., and Wallace, M. T. (2003). Neuron-specific response characteristics predict the magnitude of multisensory integration. *J. Neurophysiol.* 90, 4022–4026.

Perrault, T. J. Jr., Vaughan, J. W., Stein, B. E., and Wallace, M. T. (2005). Superior colliculus neurons use distinct operational modes in the integration of multisensory stimuli. *J. Neurophysiol.* 93, 2575–2586.

Pouget, A., Deneve, S., and Duhamel, J. R. (2002). A computational perspective on the neural basis of multisensory spatial representations. *Nat. Rev. Neurosci.* 3, 741–747.

Pouget, A., and Sejnowski, J. T. (1995). Spatial representations in the parietal cortex may use basis functions. *Adv. Neural Inf. Proc.* 7, 157–164.

Pouget, A., and Sejnowski, J. T. (1997). Spatial transformations in the parietal cortex using basis functions. *J. Cogn. Neurosci.* 9, 222–237.

Pouget, A., and Sejnowski, J. T. (2001). Simulating a lesion in a basis function model of spatial representations: comparison with hemineglect. *Psychol. Rev.* 108, 653–673.

Rizzolatti, G., Fadiga, L., Fogassi, L., and Gallese, V. (1997). The space around us. *Science* 277, 190–191.

Rizzolatti, G., Scandolara, C., Matelli, M., and Gentilucci, M. (1981). Afferent properties of periarcuate neurons in macaque monkeys. II. Visual responses. *Behav. Brain Res.* 2, 147–163.

Rolls, E. T., and Treves, A. (1998). *Neural Networks and Brain Functions.* Oxford, NY: Oxford University Press.

Rowland, B. A., Stanford, T. R., and Stein, B. E. (2007). A model of the neural mechanisms underlying multisensory integration in the superior colliculus. *Perception* 36, 1431–1443.

Sarri, M., Blankenburg, F., and Driver, J. (2006). Neural correlates of cross-modal visual-tactile extinction and of tactile awareness revealed by fMRI in a right-hemisphere stroke patient. *Neuropsychologia* 44, 2398–2410.

Schroeder, C. E., and Foxe, J. (2005). Multisensory contributions to low-level, 'unisensory' processing. *Curr. Opin. Neurobiol.* 15, 454–458.

Serino, A., Bassolino, M., Farnè, A., and Ladavas, E. (2007). Extended multisensory space in blind cane users. *Psychol. Sci.* 18, 642–648.

Stanford, T. R., Quessy, S., and Stein, B. E. (2005). Evaluating the operations underlying multisensory integration in the cat superior colliculus. *J. Neurosci.* 25, 6499–6508.

Stein, B. E., and Meredith, M. A. (1993). *The Merging of the Senses.* Cambridge, MA: MIT Press.

Stein, B. E., Stanford, T. R., Ramachandran, R., Perrault, T. J. Jr., and Rowland, B. A. (2009). Challenges in quantifying multisensory integration: alternative criteria, models, and inverse effectiveness. *Exp. Brain Res.* 198, 113–126.

Ursino, M., Cuppini, C., Magosso, E., Serino, A., and di Pellegrino, G. (2009). Multisensory integration in the superior colliculus: a neural network model. *J. Comput. Neurosci.* 26, 55–73.

Wallace, M. T., Meredith, M. A., and Stein, B. E. (1993). Converging influences from visual, auditory, and somatosensory cortices onto output neurons of the superior colliculus. *J. Neurophysiol.* 69, 1797–1809.

Wallace, M. T., Meredith, M. A., and Stein, B. E. (1998). Multisensory integration in the superior colliculus of the alert cat. *J. Neurophysiol.* 80, 1006–1010.

Wallace, M. T., Perrault, T. J. Jr., Hairston, W. D., and Stein, B. E. (2004). Visual experience is necessary for the development of multisensory integration. *J. Neurosci.* 24, 9580–9584.

Wallace, M. T., and Stein, B. E. (1994). Cross-modal synthesis in the midbrain depends on input from cortex. *J. Neurophysiol.* 71, 429–432.

Wallace, M. T., and Stein, B. E. (1997). Development of multisensory neurons and multisensory integration in cat superior colliculus. *J. Neurosci.* 17, 2429–2444.

Wallace, M. T., and Stein, B. E. (2000). Onset of cross-modal synthesis in the neonatal superior colliculus is gated by the development of cortical influences. *J. Neurophysiol.* 83, 3578–3582.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 29 November 2010; accepted: 12 April 2011; published online: 02 May 2011.*

*Citation: Cuppini C, Magosso E and Ursino M (2011) Organization, maturation and plasticity of multisensory integration: insights from computational modeling studies. Front. Psychology* ***2****:77. doi: 10.3389/fpsyg.2011.00077*

*This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.*

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 08 September 2011
doi: 10.3389/fpsyg.2011.00215

# Auditory stimulus timing influences perceived duration of co-occurring visual stimuli

***Vincenzo Romei*[1,2]*, *Benjamin De Haas*[1,2], *Robert M. Mok*[2,3] *and Jon Driver*[1,2]**

[1] *Wellcome Trust Centre for Neuroimaging at UCL, University College London, London, UK*
[2] *UCL Institute of Cognitive Neuroscience, University College London, London, UK*
[3] *Department of Psychological Sciences, Birkbeck, University of London, London, UK*

***Edited by:***
*Nadia Bolognini, University Milano Bicocca, Italy*

***Reviewed by:***
*Simon Baumann, Newcastle University, UK*
*Raymond V. Ee, University Utrecht, Netherlands*

****Correspondence:***
*Vincenzo Romei, Wellcome Trust Centre for Neuroimaging at UCL, University College London, 12 Queen Square, WC1N 3BG London, UK.*
*e-mail: v.romei@ucl.ac.uk*

There is increasing interest in multisensory influences upon sensory-specific judgments, such as when auditory stimuli affect visual perception. Here we studied whether the duration of an auditory event can objectively affect the perceived duration of a co-occurring visual event. On each trial, participants were presented with a pair of successive flashes and had to judge whether the first or second was longer. Two beeps were presented with the flashes. The order of short and long stimuli could be the same across audition and vision (audio–visual congruent) or reversed, so that the longer flash was accompanied by the shorter beep and vice versa (audio–visual incongruent); or the two beeps could have the same duration as each other. Beeps and flashes could onset synchronously or asynchronously. In a further control experiment, the beep durations were much longer (tripled) than the flashes. Results showed that visual duration discrimination sensitivity ($d'$) was significantly higher for congruent (and significantly lower for incongruent) audio–visual synchronous combinations, relative to the visual-only presentation. This effect was abolished when auditory and visual stimuli were presented asynchronously, or when sound durations tripled those of flashes. We conclude that the temporal properties of co-occurring auditory stimuli influence the perceived duration of visual stimuli and that this can reflect genuine changes in visual sensitivity rather than mere response bias.

**Keywords: multisensory integration, crossmodal interactions, response bias, signal-detection theory, audition, vision, time perception**

## INTRODUCTION

There is increasing interest within multisensory research in how auditory information can influence visual perception (e.g., Shams et al., 2000; Vroomen and de Gelder, 2000; Calvert et al., 2004; Spence and Driver, 2004; Vroomen et al., 2004; Cappe et al., 2009; Vroomen and Keetels, 2009; Leo et al., 2011). Time is one fundamental dimension for all sensory modalities, and there are now several studies that demonstrate that manipulating the temporal dimension in one modality affect perception for other modalities (e.g., Eagleman, 2008; Freeman and Driver, 2008). Perception of an event's duration can deviate from its physical characteristics (Eagleman, 2008) and in multisensory cases might be most influenced by the sensory modality that carries the most reliable temporal information (see Welch and Warren, 1980; Walker and Scott, 1981; Recanzone, 2003; Wada et al., 2003; Alais and Burr, 2004; Ernst and Bülthoff, 2004; Witten and Knudsen, 2005; Burr and Alais, 2006), as we studied here for audio–visual cases.

While the visual system typically has a higher spatial resolution than the auditory system (e.g., Witten and Knudsen, 2005) audition is usually more reliable for temporal aspects of perception (Repp and Penel, 2002; Bertelson and Aschersleben, 2003; Morein-Zamir et al., 2003; Guttman et al., 2005; Getzmann, 2007; Freeman and Driver, 2008). Accordingly vision can dominate audition in determining spatial percepts, as in the classic "ventriloquist effect" (Howard and Templeton, 1966; Thurlow and Jack, 1973; Bertelson and Radeau, 1981). Conversely, audition may dominate vision in the temporal domain (Welch and Warren, 1980; Repp and Penel, 2002; Bertelson and Aschersleben, 2003; Morein-Zamir et al., 2003; Guttman et al., 2005; Getzmann, 2007; Freeman and Driver, 2008; Kanai et al., 2011) leading to so-called "temporal ventriloquism" (e.g., Gehard and Mowbray, 1959; Bertelson and Aschersleben, 2003). Freeman and Driver (2008) found that timing of a static sound can strongly influence spatio-temporal processing of concurrent visual apparent motion. Shams et al. (2000) found that illusory percepts of multiple flashes can be induced when a single flash is accompanied by a sequence of multiple beeps. Shipley (1964) showed that changes in the physical flutter rate of a clicking sound induce simultaneous changes in the apparent flicker rate of a flashing light.

Several crossmodal effects on subjective time perception in particular have been described (e.g., Walker and Scott, 1981; Donovan et al., 2004; Chen and Yeh, 2009; Klink et al., 2011). Chen and Yeh (2009) reported in an oddball paradigm that auditory stimuli can apparently extend reported visual duration, while visual stimuli had no such impact on reported auditory duration (see also Donovan et al., 2004; Klink et al., 2011; but see also van Wassenhove et al., 2008; Aaen-Stockdale et al., 2011 for alternative accounts). But despite such suggestions of auditory influences on visual duration perception, to date it has typically been hard to establish whether such influences reflect response biases or instead genuine changes in visual sensitivity, in signal-detection terms (Macmillan and Creelman, 1991).

We sought to address this issue directly here. On each trial subjects were presented with two visual stimuli in succession (see **Figure 1**) and had to make a force-choice about which was longer, which had an objectively correct answer. The multisensory manipulation was that we could also present two sounds on each trial. In Experiment 1a each sound was presented simultaneously with a flash (see **Figure 1**). These were possible two durations for the flashes, both used on every single trial so that one flash (either the first or the second) was longer. The beep durations on a given trial could potentially agree with those for the two successive flashes (congruent condition). Alternatively the two sounds could have the reverse order of durations (incongruent condition), or else the same duration as each other (both sounds short, or both sounds long). We measured whether manipulating the auditory durations had an impact on objective performance in the visual duration discrimination task, analyzing this in terms of signal-detection theory.

To determine whether any influence of auditory durations on visual duration judgments depended on synchrony between the multisensory events, in a control study (Experiment 1b) we misaligned the onsets of auditory and visual events (by 500 ms) to produce asynchronous control conditions. If the impact of auditory durations on visual duration perceptions reflects multisensory binding (e.g., see Meredith et al., 1987; Colonius and Diederich, 2011), it should be eliminated or reduced in the asynchronous condition; whereas if instead it were simply to reflect a response bias (similar to the response that a blind observer might give when asked to report visual durations when only hearing sounds) the auditory influence should remain the same even in the new asynchronous case of Experiment 1b. Finally, in a further control situation (Experiment 2) we tripled the length of auditory durations relative to visual durations, reasoning that if auditory and visual durations mismatch sufficiently, there should be less genuine perceptual binding between them (whereas once again, the response bias of a strictly blind observer who only hears sounds, misreporting them as if they were seen, should remain the same).

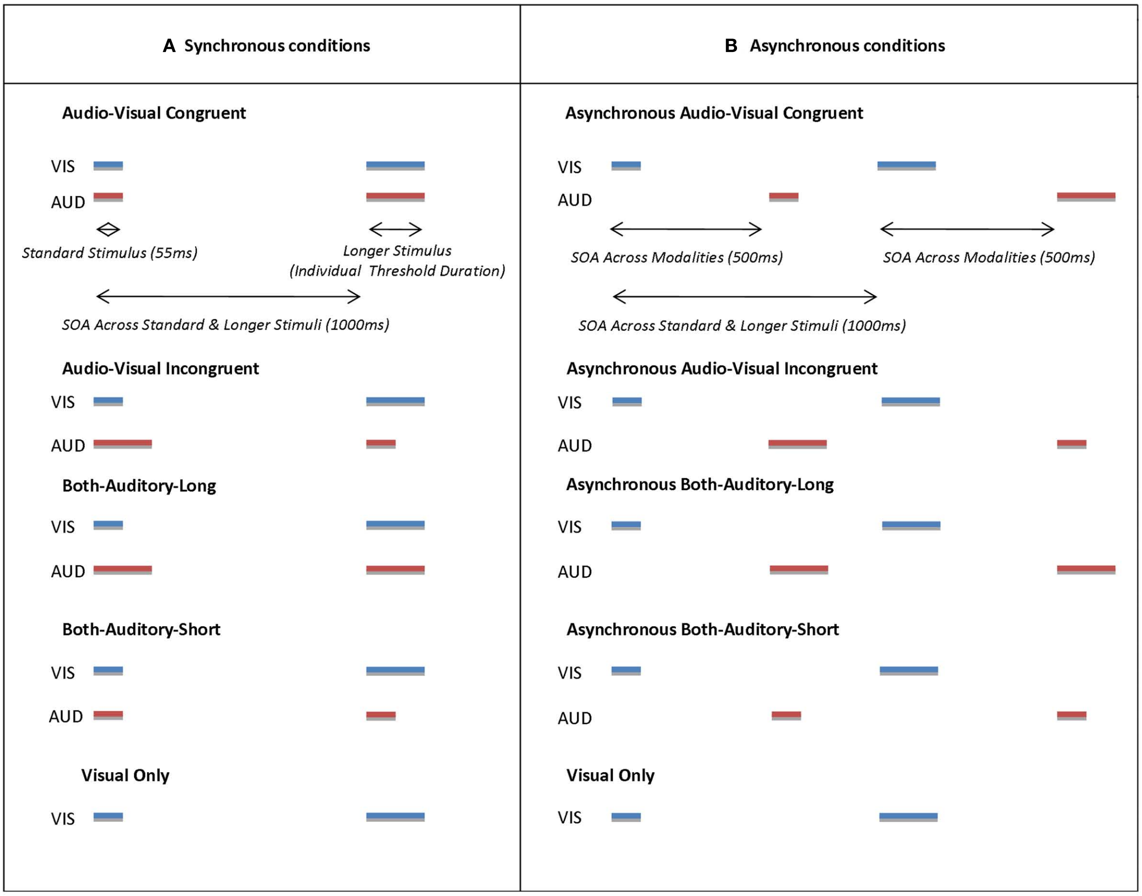

**FIGURE 1 | Schematic timelines representing conditions in Experiment 1A and 1B.** In Experiment 1A all corresponding auditory (red lines marked with label "AUD") and visual (blue lines marked with label "VIS") stimuli had the same onset (synchronous conditions) while in Experiment 1B visual stimuli always preceded auditory stimuli by 500 ms (asynchronous conditions). In both situations, each pair of stimuli within one modality was separated by a 1000-ms interval; the order of short and long stimuli could be the same across auditory and visual modalities (congruent) or reversed between them (incongruent). In the both-auditor-short condition, both successive sounds had the shorter visual duration (and vice versa for both-auditory-long). Finally a visual-only condition served as a baseline measure.

## EXPERIMENT 1

Seventeen participants with a mean age of 26.29 (range 19–35) took part in the first experiment, nine female, one left-handed. All reported normal or corrected visual acuity and normal hearing. All gave written informed consent in accord with University College London ethics approval, were naïve to the purpose of the study and were paid for their time. One participant was excluded because she showed an inconsistent pattern in the visual titration (see below). Two more were excluded because their performance in the visual-only condition was at ceiling in the main experiments, leaving 14 participants in the sample.

### APPARATUS

Stimuli were presented on a 21′ CRT display (Sony GDM-F520) in a darkened room. Participants sat with their head in a chin rest at 65 cm viewing distance. Video resolution was 1600 × 1200, with screen refresh rate of 85 Hz. Two small stereo PC speakers were placed just in front of the monitor immediately on either side of it. Stimulus control and data recording were implemented on a standard PC, running E-Prime 2 Professional (Psychology Software tools, Inc., www.pstnet.com). Unspeeded manual two-choice responses were made using a standard PC keyboard.

### STIMULI

Each visual stimulus comprised a white disk extending 1.2° in visual angle with its midpoint at 3° below a central fixation cross on a gray background. On each trial a pair of disks was flashing consecutively with varying durations from 55 to 165 ms.

The auditory stimulus was a 900-Hz pure tone sampled at 44.100 kHz with durations also varying from 55 to 165 ms. Sound level was measured with an audiometer and set to ∼70 dB(A).

### PROCEDURE

#### *Visual titration*

Only visual stimuli were presented during this part of the experiment. On each trial participants were presented with a pair of disks flashing in two consecutive time windows separated by an SOA of 1000 ms. While one of the two visual stimuli had a constant standard duration of 55 ms, the other was slightly longer, with its duration varying between 66 and 165 ms (10 possible incremental steps of one frame at 85 Hz, i.e., ∼11 ms). The latter stimulus type will be referred to as the "longer" stimulus. Each of the resulting 10 pairings of standard and longer stimuli was repeated 10 times per block. Each participant completed two to three blocks. The pairwise order of standard and longer stimuli was counterbalanced between trials, with standard-longer or longer-standard pairwise sequences being equiprobable.

On each trial participants were instructed to indicate whether the first or the second flash lasted longer, by pressing a corresponding button on the keyboard ("1" or "2"). This allowed us to identify the visual duration discrimination threshold for each participant individually. Threshold was defined as corresponding to the increase in duration for the longer stimulus whose duration allowed correct identification of it as longer in ∼75% of cases. As it turned out, for the selected threshold stimulus participants were able to discriminate differences of durations correctly in 73.78 (±1.6 SE)% of cases for those stimulus pairings containing the longer stimulus that was identified as threshold. The average duration of the longer stimuli identified as threshold was 103.4 (±3.97 SE) ms duration, for the visual disks used.

### MAIN EXPERIMENT 1

In each trial of the main experiment, participants were presented with the pair of visual stimuli previously identified as around threshold from the titration task. Again, the order of standard and longer visual stimulus was counterbalanced and equiprobable, with participants again asked to indicate which of the two consecutive flashes lasted longer. But the main experiment now consisted of 10 conditions (5 in Experiment 1a, and 5 in Experiment 1b, with these 10 all intermingled but presented separately here for ease of exposition). These 10 conditions differed with regard to whether, when, and how any sounds were presented with the flashes. Participants were emphatically instructed to ignore all sounds played during the experiment and to judge only the duration of the visual stimuli.

Two pure tone durations were selected for each participant – one lasting 55 ms and thus matching the standard visual stimulus in duration, the other auditory duration matching the participant-specific longer visual stimulus identified as threshold during the preceding visual titration task. These two pure tones were then combined with the flashes according to condition. There were two main classes of conditions: potentially synchronous (Experiment 1a) or asynchronous (Experiment 1b). In the potentially synchronous conditions, tone onset was temporally aligned with the visual onsets; whereas in the potentially asynchronous conditions, the onset of tones was delayed for 500 ms (thus 180° out of "phase" if one considers the pair of visual stimuli as a cycle, for which 180° yields the maximum possible phase offset) relative to flash onsets. In either of the potentially synchronous or asynchronous situations, there were five possible conditions: audio–visual congruent (same order of durations in the two modalities), audio–visual incongruent (opposite orders of durations in the two modalities), both-long auditory stimuli, both-short auditory stimuli, or a purely visual condition (c.f. **Figure 1**). The purely visual condition was of course actually equivalent for "synchronous" and "asynchronous" conditions, corresponding to the same condition arbitrarily divided into two separate datasets (random halves of the visual-only trials per participant), so as to a 5 × 2 factorial analysis of variance (ANOVA) on the data; see below.

Each block contained 10 repetitions for each of the 10 conditions in a randomized order. Every participant repeated three to four of these blocks.

### DATA ANALYSIS

For each participant we computed visual sensitivity ($d'$) and criterion (c) for the duration discrimination task, for each condition, using standard formulae as in Macmillan and Creelman (1997), namely:

$$d' = z(H) - z(F)$$

and

$$c = -[z(H) + z(F)]/2$$

where z(H) stands for the z-transform of the hit rate, while z(F) stands for the z-transform of the false-alarm rate. For any cases in which false-alarm rates were zero, we followed the conservative convention (as recommended by Snodgrass and Corwin, 1988; Macmillan and Creelman, 1991; c.f. Sarri et al., 2006) of adding a count of 0.5 to all cells within a single analysis.

$d'$ and criterion were analyzed using repeated-measures ANOVA, with SYNCHRONY (synchronous/asynchronous) and audio–visual CONDITION (Congruent; Incongruent; Short sounds; Long sounds, Purely visual) as within-subjects factors; followed up by pairwise $t$-tests where appropriate.

## RESULTS

The sensitivity ($d'$) results are shown in **Figure 2A** (synchronous conditions) and **Figure 2B** (asynchronous conditions), as group means with SE. Recall that synchronous/asynchronous is a dummy factor solely for the visual-only condition, which was split randomly into two separate datasets. Note the higher sensitivity specifically in the synchronous audio–visual congruent condition (**Figure 2A**, leftmost bar). The overall 5 × 2 ANOVA showed a main effect of SYNCHRONY [$F(1,13) = 12.09, p < 0.01$], a significant main effect of audio–visual CONDITION [$F(4,52) = 6.68$, $p < 0.001$] and critically a significant interaction between these two factors [$F(4,52) = 5.96, p < 0.001$].

To identify the source of the interaction, first two separate one-way ANOVAs were performed for synchronous or asynchronous datasets, with the five-level factor of condition. While none of asynchronous conditions differed from each other [$F(4,52) = 0.56$, $p = 0.69$ for the main effect], the synchronous conditions did [$F(4,52) = 10.91$, $p < 0.00001$]. Exploratory pairwise $t$-tests for the asynchronous conditions confirmed no significant differences between any (all $p > 0.20$). Pairwise $t$-tests for the synchronous conditions showed that sensitivity in the synchronous audio–visual congruent condition ($d' = 1.93 \pm 0.23$ SE) was significantly higher than in all the other conditions, as follows: (i) versus the synchronous audio–visual incongruent condition [$d' = 0.20 \pm 0.15$ SE; $t(13) = 4.93$; $p < 0.001$]; (ii) versus the both-auditory-short condition [$d' = 1.07 \pm 0.16$ SE; $t(13) = 3.63$; $p < 0.01$]; (iii) versus the both-auditory-long condition [$d' = 1.11 \pm 0.19$ SE; $t(13) = 3.0, p = 0.01$]. The trend for the somewhat lower $d'$ overall in the asynchronous than synchronous experiment did not approach significance.

When compared to the visual-only baseline measure ($d' = 1.01 \pm 0.18$ SE), we found that: (i) visual duration discrimination was significantly enhanced in the synchronous audio–visual congruent condition [$t(13) = -3.38$; $p < 0.01$]; (ii) was significantly impaired in the synchronous audio–visual incongruent condition [$t(13) = 3.44$, $p < 0.01$]; (iii) was not significantly affected in the both-auditory-long or short conditions (all $p > 0.71$).

A comparable two-way ANOVA on criterion scores instead revealed no significant results [main effect of SYNCHRONY ($F(1,13) = 4.17$, $p = 0.07$); main effect of audio–visual CONDITION ($F(4,52) = 0.35$, $p = 0.83$); interaction between the two ($F(4,52) = 1.10$, $p = 0.36$); all n.s.].

## DISCUSSION

In Experiment 1 we found that objective duration discrimination for visual stimuli was objectively modulated by the duration of co-occurring auditory stimuli, but only when those auditory stimuli were synchronous with the visual events, rather than being delayed by 500 ms. Specifically, visual duration discrimination sensitivity ($d'$) was enhanced for congruent-duration, synchronous auditory stimuli but decreased for incongruent-duration, synchronous auditory stimuli. Neither a sensitivity enhancement nor a sensitivity decrease could be observed for the asynchronous conditions, in which the sounds were now delayed by 500 ms so as to be (maximally) out-of-phase with the flashes. This elimination of the effect for the asynchronous case, together with the observed impact on $d'$ (rather than criterion) for the synchronous cases, indicates a

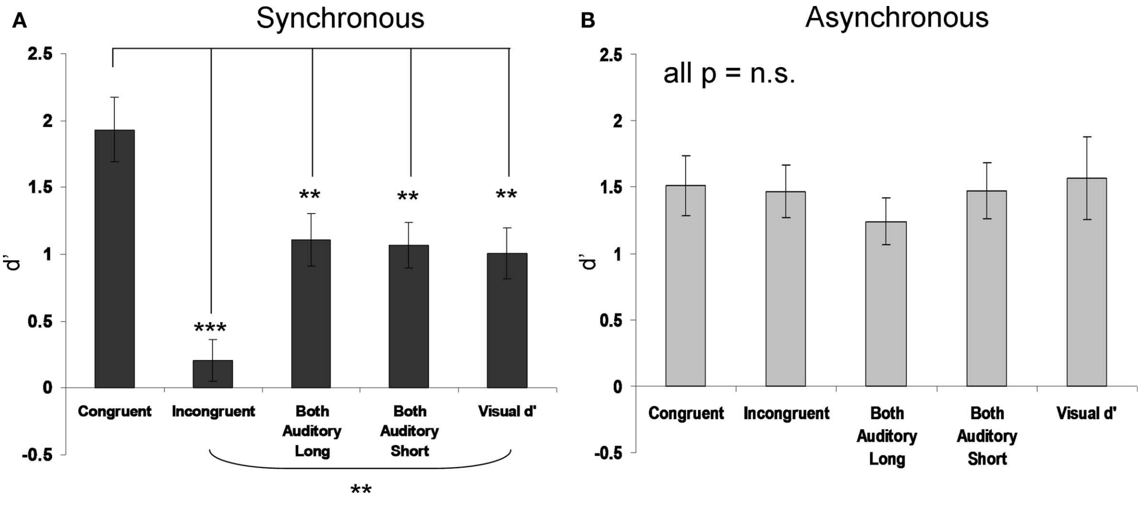

**FIGURE 2 | Mean visual duration discrimination sensitivity (d′), SEM indicated) for each condition in Experiment 1a (A) and 1b (B).** Asterisks in 1a indicate significant differences relative to the synchronous–congruent condition that gave best performance (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$), see leftmost bar in left graph. None of the pairwise contrasts were significant for the asynchronous conditions (see right graph).

genuine multisensory impact on visual perception, rather than a mere response bias as might be evident if a blind observer had to guess the response based solely on the sounds.

We suggest that the strong effect of auditory duration on perception of synchronous visual events reflects crossmodal binding between them (e.g., see also Meredith et al., 1987; Vroomen and Keetels, 2010; Colonius and Diederich, 2011), plus weighting of the auditory duration when bound due to the higher precision of temporal coding for audition than vision (see Introduction). A visual event is evidently perceived as longer when co-occurring with a slightly longer auditory event that is parsed as part of the same, single multisensory event. But we reasoned that if this is indeed a genuine perceptual effect as we document, rather than merely a response bias, then there should presumably be a limit to how far a longer auditory event can "stretch" perception of a visual event. If the auditory event were to endure much, much longer than the corresponding visual event, it should become less plausible that they arise from the same single external crossmodal event, and the auditory influence should begin to wane. We tested this in Experiment 2.

## EXPERIMENT 2

This study repeated the critical conditions where effects had been apparent in Experiment 1 (i.e., synchronous–congruent, synchronous–incongruent, plus visual-only baseline) except that now either the sounds had the same possible durations as the flashes, or else the sounds were tripled in duration such that they would no longer plausibly correspond to the flashes. A mere response bias, akin to a blind observer simply reporting the durations of the sounds, should lead to a similar outcome in either case; whereas the genuine perceptual effect we have documented should reduce (or even potentially disappear) in the mismatching tripled situation.

### METHODS

#### *Participants*

Fifteen new participants with a mean age of 24.53 (range 18–34) took part, seven females, one left-handed, all reporting normal or corrected visual acuity, and normal hearing.

#### *Apparatus and stimuli*

The setup was as for Experiment 1, but with less conditions and two new ones. We repeated the synchronous–congruent, synchronous–incongruent, and visual-only conditions. We also added two new conditions that were as for the synchronous–congruent and synchronous–incongruent except with tripled auditory durations. In these two new tripled conditions, each sound still onset simultaneously with each flash, but the sounds now lasted three times as long, so that with their much later offset they did not match the visual events so well.

### PROCEDURE

#### *Visual titration*

This aspect of the procedure was the same as in Experiment 1. On average, participants were able to discriminate durations correctly for 76.33 (±1.5 SE)% of cases for those stimulus pairings that contained the longer visual stimulus identified as threshold. The average duration of the longer visual stimuli identified as threshold was 94.5 (±3.49 SE) ms duration, so 49.5 ms longer than the standard stimulus, with the visual disks used.

### MAIN EXPERIMENT

The procedure resembled Experiment 1, but with only five conditions, two of which were new. The purely visual baseline, synchronous–congruent, and synchronous–incongruent conditions were as before. The two new conditions were tripled-sound-synchronous–congruent and tripled-sound-synchronous–incongruent condition. These were exactly like their untripled counterpart, except that the duration of each sound was three times as long.

### DATA ANALYSIS

For each participant and condition we computed sensitivity ($d'$) and criterion (c) for each stimulus condition as in Experiment 1. $d'$ or criterion for the four audio–visual conditions were analyzed using repeated-measures two-way ANOVA, with SOUND LENGTH (tripled or untripled) and audio–visual duration CONGRUENCY (congruent versus incongruent) as factors. In addition pairwise $t$-tests compared performance in the purely visual baseline against the remaining conditions.

### RESULTS

The sensitivity ($d'$) results are shown in **Figure 3**, as group means with SE. Note the higher sensitivity in the audio–visual untripled synchronous–congruent condition (leftmost bar), replicating the effects obtained in Experiment 1a (c.f. **Figure 2A**).

The two-way ANOVA showed no main effect of SOUND LENGTH [$F(1,14) = 1.94$, $p = 0.18$], a significant main effect of CONGRUENCY [$F(1,14) = 49.33$, $p < 0.00001$] and a significant interaction between these two factors [$F(1,14) = 10.99$, $p < 0.01$], because the congruent/incongruent difference was larger in the untripled than tripled case. Sensitivity in the audio–visual untripled synchronous–congruent condition ($d' = 2.07 \pm 0.16$ SE) was significantly higher than in all the other conditions, as follows: (i) versus the visual-only duration condition [$d' = 1.58 \pm 0.16$ SE; $t(14) = 2.86$; $p = 0.013$]; (ii) versus the tripled congruent condition [$d' = 1.72 \pm 0.18$ SE; $t(14) = 2.82$; $p = 0.014$]; (iii) versus the untripled incongruent condition [$d' = 0.11 \pm 0.26$ SE; $t(14) = 8.85$; $p < 0.000001$]; (iv) versus the tripled incongruent condition [$d' = 0.82 \pm 0.25$ SE; $t(14) = 4.8$, $p < 0.001$].

Thus, by prolonging the duration of the auditory stimuli to triple that of the visual stimuli, the significant enhancement (relative to visual-only baseline) obtained for the congruent audio–visual durations was abolished. While even triple-duration auditory stimuli still produced some sensitivity decrease for incongruent stimuli, even this remaining decrease was still significantly reduced for the tripled versus untripled case [$t(14) = -2.69$, $p = 0.018$].

A comparable analysis of criterion scores instead revealed no significant results [e.g., for the two-way ANOVA, main effects of SOUND LENGTH ($F(1,14) = 3.29$, $p = 0.09$); main effect of CONGRUENCY ($F(1,14) = 0.002$, $p = 0.96$); interaction between the two ($F(1,14) = 2.15$, $p = 0.16$), all n.s.].

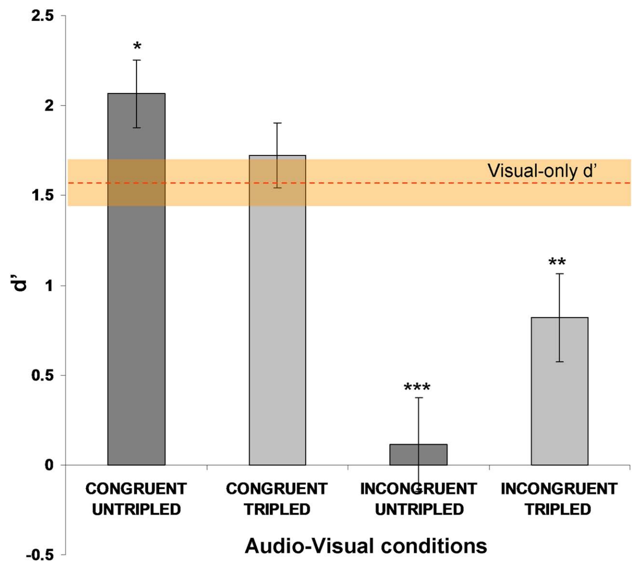

**FIGURE 3 | Mean visual duration discrimination sensitivity (d′), SEM indicated) for each condition in Experiment 2.** Asterisks above bars point to significant differences relative to the visual-only baseline, with the latter represented here by the orange dashed line with SEM shading. The significant enhancement or decrease of sensitivity, for the untripled congruent and incongruent conditions (respectively) replicates the findings of Experiment 1a. These effects were eliminated or reduced (respectively) for the corresponding two new tripled conditions, in which a sound still onset concurrently with each flash, but the sounds now were three times as long.

## DISCUSSION

We replicated the results of Experiment 1a for the shared conditions in Experiment 2, showing significant enhancement of objective visual duration discrimination sensitivity (*d*′) by congruent auditory stimuli and a significant decrease of sensitivity for incongruent stimuli. The new finding was that the sensitivity enhancement for congruent stimuli (i.e., same pairwise successive order of longer and shorter) relative to the visual-only baseline was completely abolished for prolonged auditory stimuli that onset concurrently with the visual flashes but endured three times longer. Some sensitivity decrease for incongruent audio–visual pairings remained but at a significantly reduced level in the tripled-sound-duration case. These results show that the impact of auditory durations on visual duration discrimination is larger when the sounds and flashes endure for a similar order of magnitude, being significantly diminished when the sounds endure three times longer than the flashes.

## GENERAL DISCUSSION

In two experiments we tested the influence of sound duration on objective discrimination of the duration of visual events. Our results accord with but go beyond previous reports (e.g., Walker and Scott, 1981; Donovan et al., 2004; Chen and Yeh, 2009; Klink et al., 2011) that auditory stimuli can impact significantly on judgments for the duration of co-occurring visual stimuli. The new findings extend previous work by showing that: (1) not only can incongruent auditory stimuli significantly impair objective visual performance, but congruent auditory stimuli can benefit visual duration judgments; (2) this applies for visual sensitivity (*d*′) to visual duration in signal-detection terms, rather than affecting mere response bias or criterion; (3) this impact of auditory duration on perception of visual duration depends on whether the audio–visual onsets are synchronous, being eliminated when the sounds lagged here; (4) it also depends on whether the auditory and visual events are similar in length, being reduced, or eliminated when the sounds are triple the duration of corresponding visual events.

A previous study by Donovan et al. (2004) used a similar approach to ours. They investigated the influence of task-irrelevant auditory information on a visual task, but with participants judging whether two sequential visual events were presented for the same or different lengths of time. They found a similar trend to the present pattern for congruent versus incongruent audio–visual conditions, and their effect was abolished for asynchronous conditions, reminiscent of the present Experiment 1b. But Donovan et al. (2004) did not titrate the duration of visual stimuli to a threshold level; did not present a visual-only condition to provide a baseline for assessing any multisensory benefit or cost; did not calculate signal-detection scores.

More recently, Klink et al. (2011) extensively tested crossmodal influences on duration perception, confirming auditory over visual dominance for time perception. These authors adopted a duration discrimination task, similar to the one presented here. In line with our results, they found a reduction in visual duration discrimination accuracy with incongruent auditory stimuli. But they did not test the effect of audio–visual congruent stimuli, and their control for possible response biases was very different to our own (a grouping experiment, Experiment 5 in Klink et al., 2011). Our findings accord with such prior work (see also Introduction) in showing a clear influence of audition on visual duration judgments. We show in signal-detection terms that visual *d*′ for duration discrimination can not only be impaired by incongruent auditory timing information, but for the first time, also significantly enhanced by congruent auditory information; plus we document some of the boundary conditions for this [in terms not only of synchronous onset across the modalities (see Experiment 1 here) but also in terms of fairly well-matching duration scale (see Experiment 2 here)].

How could such auditory influences over visual duration perception arise? An extensive literature on the many possible mechanisms for time perception has built up (for recent reviews, see Ivry and Schlerf, 2008; Grondin, 2010). One long tradition in the field of time perception posits the presence of central mechanisms for time estimation, such as an internal clock or clocks (e.g., Ivry and Richardson, 2002; Rosenbaum, 2002; Schöner, 2002; Wing, 2002). Some research suggests that such central-clock(s) might operate supramodally. In apparent support of this view are findings that several brain areas (e.g., see Ivry and Keele, 1989; Harrington et al., 1998; Rao et al., 2001; Leon and Shadlen, 2003; Coull et al., 2004; Bueti et al., 2008) are implicated in estimation and representation of time independently from the sensory modality of the stimuli in question, although it should be noted that some of the time judgments used were on longer scales than here.

Other authors have argued that there may be no need to invoke internal "clocks" to describe some timing behaviors (e.g., Zeiler, 1999; Jantzen et al., 2005; Karmarkar and Buonomano, 2007; Ivry and Schlerf, 2008). Recent findings in the field of visual perception, for example, have led to the development of seemingly more modality-specific perspectives (e.g., Yarrow et al., 2001; Ghose and Maunsell, 2002; Morrone, 2005; Johnston et al., 2006; Shuler and Bear, 2006; Xuan et al., 2007), suggesting that estimates for the duration of visual signals could be embedded within the visual modality itself.

Here we showed that visual duration sensitivity can be significantly impaired or enhanced by auditory stimuli that are likely to be parsed as reflecting the same external "event" as the affected visual event. The central-clock perspective might consider this to arise at some internal timing process that is shared between modalities. On the other hand, our results might also be reconciled with visual duration judgments arising within visual cortex itself, provided it is acknowledged that auditory can also impact upon visual cortex (for which an increasing body of evidence now exists; e.g., Martuzzi et al., 2007; Romei et al., 2007, 2009; Wang et al., 2008; Bolognini et al., 2010; Cappe et al., 2010; Noesselt et al., 2010; Bolognini and Maravita, 2011; c.f. Ghazanfar and Schroeder, 2006; Driver and Noesselt, 2008 for extensive review). It would be useful to combine the present behavioral paradigm with neural measures in future work; and also to study the impact of neural disruptions, such as transcranial magnetic stimulation (TMS), targeting visual cortex or auditory cortex or heteromodal cortex (see Kanai et al., 2011). Other future extensions of our paradigm could investigate whether lagging auditory events by different parametric amounts in the asynchronous condition would lead to a graded or categorical change in results; and the possible impact of introducing asynchrony by making auditory events lead instead. Here we had lagged the sounds in our asynchronous condition by a full 500 ms, in order to generate the maximum 180° shift to visual and auditory events being "out-of-phase" in terms of the cycle we used.

A further interesting question for future extensions of our paradigm concerns the possible role of attention in the multisensory effect on sensitivity that we have identified. Indeed there is now a growing literature on the possible role of attention in multisensory integration (see Sanabria et al., 2007; Talsma et al., 2010). Studies of some multisensory phenomena suggesting no role for attention (e.g., Bertelson et al., 2000); while others on different multisensory phenomena suggest a key attentional role (e.g., van Ee et al., 2009). Accordingly it is an empirical issue whether the new multisensory phenomena that we have uncovered may depend on attention or not. The present results already make clear that audio–visual integration can genuinely affect visual sensitivity ($d'$) here, and that this depends on audio–visual synchrony.

Finally, given the evidently perceptual nature of the objective improvements in visual duration discrimination that we observed here due to appropriately timed sounds, it would be intriguing to study whether a slightly longer sound paired with a concurrent visual event can not only extend the apparent duration of that visual event, but actually improve visual perception of its *other* (non-temporal) visual qualities, in a similar manner to the visual improvement found for a genuinely longer visual stimulus (see Berger et al., 2003).

## ACKNOWLEDGMENTS

Vincenzo Romei, Benjamin De Haas, and Jon Driver were supported by the Wellcome Trust. Jon Driver is a Royal Society Anniversary Research Professor.

## REFERENCES

Aaen-Stockdale, C., Hotchkiss, J., Heron, J., and Whitaker, D. (2011). Perceived time is spatial frequency dependent. *Vision Res.* 51, 1232–1238.

Alais, D., and Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.* 14, 257–262.

Berger, T. D., Martelli, M., and Pelli, D. G. (2003). Flicker flutter: is an illusory event as good as the real thing? *J. Vis.* 3, 406–412.

Bertelson, P., and Aschersleben, G. (2003). Temporal ventriloquism: crossmodal interaction on the time dimension: 1. evidence from auditory – visual temporal order judgment. *Int. J. Psychophysiol.* 50, 147–155.

Bertelson, P., and Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Percept. Psychophys.* 29, 578–584.

Bertelson, P., Vroomen, J., de Gelder, B., and Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Percept. Psychophys.* 62, 321–332.

Bolognini, N., and Maravita, A. (2011). Uncovering multisensory processing through non-invasive brain stimulation. *Front. Psychol.* 2:46. doi: 10.3389/fpsyg.2011.00046

Bolognini, N., Senna, I., Maravita, A., Pascual-Leone, A., and Merabet, L. B. (2010). Auditory enhancement of visual phosphene perception: the effect of temporal and spatial factors and of stimulus intensity. *Neurosci. Lett.* 477, 109–114.

Bueti, D., Bahrami, B., and Walsh, V. (2008). The sensory and association cortex in time perception. *J. Cogn. Neurosci.* 20, 1–9.

Burr, D., and Alais, D. (2006). Combining visual and auditory information. *Prog. Brain Res.* 155, 243–258.

Calvert, G. A., Spence, C., and Stein, B. E. (eds). (2004). *The Handbook of Multisensory Processes.* Cambridge: MIT Press.

Cappe, C., Thut, G., Romei, V., and Murray, M. M. (2009). Selective integration of auditory-visual looming cues by humans. *Neuropsychologia* 47, 1045–1052.

Cappe, C., Thut, G., Romei, V., and Murray, M. M. (2010). Auditory-visual multisensory interactions in humans: timing, topography, directionality, and sources. *J. Neurosci.* 30, 12572–12580.

Chen, K., and Yeh, S. (2009). Asymmetric cross-modal effects in time perception. *Acta Psychol. (Amst.)* 130, 225–234.

Colonius, H., and Diederich, A. (2011). Computing an optimal time window of audiovisual integration in focused attention tasks: illustrated by studies on effect of age and prior knowledge. *Exp. Brain Res.* 212, 327–337.

Coull, J. T., Vidal, F., Nazarian, B., and Macar, F. (2004). Functional anatomy of the attentional modulation of time estimation. *Science* 303, 1506–1508.

Donovan, C. L., Lindsay, D. S., and Kingstone, A. (2004). Flexible and abstract resolutions to crossmodal conflicts. *Brain Cogn.* 56, 1–4.

Driver, J., and Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on "sensory-specific" brain regions, neural responses, and judgments. *Neuron* 57, 11–23.

Eagleman, D. M. (2008). Human time perception and its illusions. *Curr. Opin. Neurobiol.* 18, 131–136.

Ernst, M. O., and Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends Cogn. Sci. (Regul. Ed.)* 8, 162–169.

Freeman, E., and Driver, J. (2008). Direction of visual apparent motion driven solely by timing of a static sound. *Curr. Biol.* 18, 1262–1266.

Gehard, J. W., and Mowbray, G. H. (1959). On discriminating the rate of visual flicker and auditory flutter. *Am. J. Psychol.* 72, 521–529.

Getzmann, S. (2007). The effect of brief auditory stimuli on visual apparent motion. *Perception* 36, 1089–1103.

Ghazanfar, A. A., and Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends Cogn. Sci. (Regul. Ed.)* 10, 278–285.

Ghose, G. M., and Maunsell, J. H. R. (2002). Attentional modulation in visual cortex depends on task timing. *Nature* 419, 616–620.

Grondin, S. (2010). Timing and time perception: a review of recent behavioral and neuroscience findings and theoretical directions. *Atten. Percept. Psychophys.* 72, 561–582.

Guttman, S. E., Gilroy, L. A., and Blake, R. (2005). Hearing what the eyes see: auditory encoding of visual temporal sequences. *Psychol. Sci.* 16, 228–235.

Harrington, D. L., Haaland, K. Y., and Hermanowicz, N. (1998). Temporal processing in the basal ganglia. *Neuropsychology* 12, 3–12.

Howard, I. P., and Templeton, W. B. (1966). *Human Spatial Orientation.* London: Wiley.

Ivry, R., and Keele, S. (1989). Timing functions of the cerebellum. *J. Cogn. Neurosci.* 1, 36–152.

Ivry, R. B., and Richardson, T. C. (2002). Temporal control and coordination: the multiple timer model. *Brain Cogn.* 48, 117–132.

Ivry, R. B., and Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends Cogn. Sci. (Regul. Ed.)* 12, 273–280.

Jantzen, K. J., Steinberg, F. L., and Kelso, J. A. S. (2005). Functional MRI reveals the existence of modality and coordination-dependent timing networks. *Neuroimage* 25, 1031–1042.

Johnston, A., Arnold, D. H., and Nishida, S. (2006). Spatially localised distortions of event time. *Curr. Biol.* 16, 472–479.

Kanai, R., Lloyd, H., Bueti, D., and Walsh, V. (2011). Modality-independent role of the primary auditory cortex in time estimation. *Exp. Brain Res.* 209, 465–471.

Karmarkar, U. R., and Buonomano, D. V. (2007). Timing in the absence of clocks: encoding time in neural network states. *Neuron* 53, 427–438.

Klink, P. C., Montijn, J. S., and van Wezel, R. J. (2011). Crossmodal duration perception involves perceptual grouping, temporal ventriloquism, and variable internal clock rates. *Atten. Percept. Psychophys.* 73, 219–236.

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

Leon, M. I., and Shadlen, M. N. (2003). Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron* 38, 317–327.

Macmillan, N. A., and Creelman, C. D. (1991). *Detection Theory: A User's Guide.* Cambridge: Cambridge University Press.

Macmillan, N. A., and Creelman, C. D. (1997). d'Plus: a program to calculate accuracy and bias measures from detection and discrimination data. *Spat. Vis.* 11, 141–143.

Martuzzi, R., Murray, M. M., Michel, C. M., Thiran, J. P., Maeder, P. P., Clarke, S., and Meuli, R. A. (2007). Multisensory interactions within human primary cortices revealed by BOLD dynamics. *Cereb. Cortex* 17, 1672–1679.

Meredith, M. A., Nemitz, J. W., and Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *J. Neurosci.* 7, 3215–3229.

Morein-Zamir, S., Soto-Faraco, S., and Kingstone, A. (2003). Auditory capture of vision: examining temporal ventriloquism. *Brain Res. Cogn. Brain Res.* 17, 154–163.

Morrone, M. C. (2005). Saccadic eye movements cause compression of time as well as space. *Nat. Neurosci.* 8, 950–954.

Noesselt, T., Tyll, S., Boehler, C. N., Budinger, E., Heinze, H. J., and Driver, J. (2010). 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. *J. Neurosci.* 30, 13609–13623.

Rao, S. M., Mayer, A. R., and Harrington, D. L. (2001). The evolution of brain activation during temporal processing. *Nat. Neurosci.* 4, 317–323.

Recanzone, G. H. (2003). Auditory influences on visual temporal rate perception. *J. Neurophysiol.* 89, 1078–1093.

Repp, B. H., and Penel, A. (2002). Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences. *J. Exp. Psychol. Hum. Percept. Perform.* 28, 1085–1099.

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

Romei, V., Murray, M. M., Merabet, L. B., and Thut, G. (2007). Occipital transcranial magnetic stimulation has opposing effects on visual and auditory stimulus detection: implications for multisensory interactions. *J. Neurosci.* 27, 11465–11472.

Rosenbaum, D. A. (2002). Time, space, and short-term memory. *Brain Cogn.* 48, 52–65.

Sanabria, D., Soto-Faraco, S., and Spence, C. (2007). Spatial attention and audiovisual interactions in apparent motion. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 927–937.

Sarri, M., Blankenburg, F., and Driver, J. (2006). Neural correlates of cross-modal visual-tactile extinction and of tactile awareness revealed by fMRI in a right-hemisphere stroke patient. *Neuropsychologia* 44, 2398–2410.

Schöner, G. (2002). Timing, clocks, and dynamical systems. *Brain Cogn.* 48, 31–51.

Shams, L., Kamitani, Y., and Shimojo, S. (2000). What you see is what you hear. *Nature* 408, 788.

Shipley, T. (1964). Auditory flutter-driving of visual flicker. *Science* 145, 1328–1330.

Shuler, M. G., and Bear, M. F. (2006). Reward timing the primary visual cortex. *Science* 311, 1606–1609.

Snodgrass, J. G., and Corwin, J. (1988). Pragmatics of measuring recognition memory: applications to dementia, and amnesia. *J. Exp. Psychol. Gen.* 117, 34–50.

Spence, C., and Driver, J. (eds). (2004). *Crossmodal Space and Crossmodal Attention.* Oxford: Oxford University Press.

Talsma, D., Senkowski, D., Soto-Faraco, S., and Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends Cogn. Sci. (Regul. Ed.)* 14, 400–410.

Thurlow, W. R., and Jack, C. E. (1973). Certain determinants of the "ventriloquism effect." *Percept. Mot. Skills* 36,1171–1184.

van Ee, R., van Boxtel, J. J., Parker, A. L., and Alais, D. (2009). Multisensory congruency as a mechanism for attentional control over perceptual selection. *J. Neurosci.* 29, 11641–11169.

van Wassenhove, V., Buonomano, D. V., Shimojo, S., and Shams, L. (2008). Distortions of subjective time perception within and across senses. *PLoS ONE* 3, e1437. doi: 10.1371/journal.pone.0001437

Vroomen, J., and de Gelder, B. (2000). Sound enhances visual perception: cross-modal effects of auditory organization on vision. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 1583–1590.

Vroomen, J., de Gelder, B., and Vroomen, J. (2004). Temporal ventriloquism: sound modulates the flash-lag effect. *J. Exp. Psychol. Hum. Percept. Perform.* 30, 513–518.

Vroomen, J., and Keetels, M. (2009). Sounds change four-dot masking. *Acta Psychol. (Amst.)* 130, 58–63.

Vroomen, J., and Keetels, M. (2010). Perception of intersensory synchrony: a tutorial review. *Atten. Percept. Psychophys.* 72, 871–884.

Wada, Y., Kitagawa, N., and Noguchi, K. (2003). Audio-visual integration in temporal perception. *Int. J. Psychophysiol.* 50, 117–124.

Walker, J. T., and Scott, K. J. (1981). Auditory – visual conflicts in the perceived duration of lights, tones and gaps. *J. Exp. Psychol. Hum. Percept. Perform.* 7, 1327–1339.

Wang, Y., Celebrini, S., Trotter, Y., and Barone, P. (2008). Visuo-auditory interactions in the primary visual cortex of the behaving monkey: electrophysiological evidence. *BMC Neurosci.* 9, 79. doi: 10.1186/1471-2202-9-79

Welch, R. B., and Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychol. Bull.* 88, 638–667.

Wing, A. M. (2002). Voluntary timing and brain function: an information processing approach. *Brain Cogn.* 48, 7–30.

Witten, I. B., and Knudsen, E. I. (2005). Why seeing is believing: merging auditory and visual worlds. *Neuron* 48, 489–496.

Xuan, B., Zhan, D., He, S., and Chen, X. (2007). Larger stimuli are judged to last longer. *J. Vis.* 7, 1–5.

Yarrow, K., Haggard, P., Heal, R., Brown, P., and Rothwell, J. C. (2001). Illusory perceptions of space and time preserve cross-saccadic perceptual continuity. *Nature* 414, 302–305.

Zeiler, M. D. (1999). Time without clocks. *J. Exp. Anal. Behav.* 71, 288–291.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 07 July 2011; accepted: 18 August 2011; published online: 08 September 2011.*

*Citation: Romei V, De Haas B, Mok RM and Driver J (2011) Auditory stimulus timing influences perceived duration of co-occurring visual stimuli. Front. Psychology* ***2****:215. doi: 10.3389/fpsyg.2011.00215*

*This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.*

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 02 January 2012
doi: 10.3389/fpsyg.2011.00391

# Auditory motion capturing ambiguous visual motion

**Arjen Alink[1,2,3]*, Felix Euler[1,2], Elena Galeano[1], Alexandra Krugliak[3,4], Wolf Singer[1,2,5,6] and Axel Kohler[1,2,7]**

[1] Department of Neurophysiology, Max Planck Institute for Brain Research, Frankfurt am Main, Germany
[2] Brain Imaging Center, Goethe University, Frankfurt am Main, Germany
[3] Cognition and Brain Sciences Unit, Medical Research Council, Cambridge, UK
[4] Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, Netherlands
[5] Frankfurt Institute for Advanced Studies, Goethe University, Frankfurt am Main, Germany
[6] Ernst Strüngmann Institute in Cooperation with Max Planck Society, Frankfurt am Main, Germany
[7] Department of Psychiatric Neurophysiology, University of Bern, Bern, Switzerland

**Edited by:**
Angelo Maravita, University of Milano Bicocca, Italy

**Reviewed by:**
Olivia Carter, University of Melbourne, Australia
Sophie Wuerger, The University of Liverpool, UK

***Correspondence:**
Arjen Alink, Cognition and Brain Sciences Unit, Medical Research Council, 15 Chaucer Road, Cambridge CB2 7EF, UK.
e-mail: arjen.alink@mrc-cbu.cam.ac.uk

In this study, it is demonstrated that moving sounds have an effect on the direction in which one sees visual stimuli move. During the main experiment sounds were presented consecutively at four speaker locations inducing left or rightward auditory apparent motion. On the path of auditory apparent motion, visual apparent motion stimuli were presented with a high degree of directional ambiguity. The main outcome of this experiment is that our participants perceived visual apparent motion stimuli that were ambiguous (equally likely to be perceived as moving left or rightward) more often as moving in the same direction than in the opposite direction of auditory apparent motion. During the control experiment we replicated this finding and found no effect of sound motion direction on eye movements. This indicates that auditory motion can capture our visual motion percept when visual motion direction is insufficiently determinate without affecting eye movements.

**Keywords: audiovisual, multisensory integration, motion capture, Bayesian, bistable, eye movement**

## INTRODUCTION

Each of our senses provides us with qualitatively different impressions about the objects and events that surround us. Frequently, events elicit impressions in more than one sensory modality. For example, the event of someone walking toward you can be perceived by seeing the person moving toward you as well by hearing the sound of footsteps getting louder. An interesting question that arises here is how an observer combines visual and auditory signals into a unified multisensory percept. In the current study, we attempted to address this question for the integration of auditory and visual motion signals.

Psychophysical studies have shown that moving visual stimuli can substantially alter the way in which we perceive auditory motion. For example, it has been demonstrated that visual motion can induce an auditory motion after-effect (Kitagawa and Ichihara, 2002) and that visual motion can capture the perceived direction of moving sounds (Soto-Faraco et al., 2002, 2003, 2004, 2005; Sanabria et al., 2007). The findings of studies investigating effects of moving sounds on perceived visual motion have been somewhat inconsistent. Several studies showed little to no effect of sound motion on perceived visual motion (Soto-Faraco et al., 2003; Alais and Burr, 2004), while Meyer and Wuerger (2001) did observe a bias in visual motion perception toward the direction of sound motion when visual motion direction was ambiguous. In addition, auditory motion has been shown to enhance the sensitivity to visual motion when auditory and visual signals share the same direction and location (Wuerger et al., 2003; Meyer et al., 2005) or when visual motion has biological properties (Brooks et al., 2007; Arrighi et al., 2009). Moreover, it has been demonstrated that moving sounds can cause one to perceive a static visual stimulus as moving (Hidaka et al., 2009, 2011; Teramoto et al., 2010) and prolong binocular-dominance periods for moving visual stimuli with a congruent motion direction (Conrad et al., 2010). None of these studies, however, has measured eye movements. Eye movements have been shown to influence how ambiguous visual and tactile motion is perceived (Laubrock et al., 2005; Carter et al., 2008). Therefore, it is possible that previously shown effects of sound motion on visual motion perception are mediated by direction specific effects of sound motion on eye movements.

The main finding of this study is that ambiguous visual apparent motion stimuli are perceived more often as moving in the same than the opposite direction as simultaneously presented auditory motion. Alike Meyer and Wuerger (2001), we observed a maximal effect of sound motion direction on visual motion perception when visual motion direction was maximally ambiguous. The findings of the main experiment were replicated by the control experiment during which we also measured eye movements of our participants. We did not observe an effect of sound motion direction on eye movements, providing first direct evidence that capture of visual motion by auditory motion cannot be explained by concomitant eye movements.

## MATERIALS AND METHODS

### PARTICIPANTS

Nineteen healthy volunteers participated in the main experiment (age range, 20–31 years; 11 females) and 10 participants took part in the control experiment (20–47 years, 8 females). All participants had normal hearing and normal or corrected-to-normal vision and gave their informed consent after being introduced to the experimental procedure in accordance with the Declaration of Helsinki.

## STIMULI AND PROCEDURE

### *Main experiment*

Participants were seated in a darkened room 2.2 m away from a white wall on which a beamer (Acer, PD322, 60 Hz) projected the visual stimulus with a width of 44.5° and a height of 33.4°. The background color of the stimulus was gray (luminance = 6.2 cd/$m^2$) and the stimulus contained a black aperture (luminance = 0.8 cd/$m^2$) in its center, which had a radius of 5.6°. On the horizontal midline of the stimulus, four black speakers (Trust 5.1 Surround Speaker Set SP-6210) were attached to the wall. The far-left speaker was located 20.7° left from the image center and the mid-left speaker 6.9° left from the image center. The far-right speaker was located 20.7° right from the image center and the mid-right speaker 6.9° right from the image center. The inter-speaker distances were 13.8° between all adjacent speakers. All images and sounds were generated using Presentation software (version 12.2, Neurobehavioral Systems, Inc.).

During each trial we presented visual apparent motion stimuli in the aperture that were similar to those employed by Williams et al. (2003). Hence, in the aperture we presented visual stimuli consisting of columns of gray disks (radius = 0.3°, luminance = 6.2 cd/$m^2$) with a fixed inter-column distance of 1.8° and a fixed inter-row distance of 2.1°. A visual apparent motion stimulus consisted of two images. The first image was presented for 200 ms with a randomized horizontal position of the disk columns and was followed by a second image that had the disk columns shifted to the right. The degree of the rightward shift varied from 8/30 of the inter-column distance to 22/30 of the inter-column distance in steps of 1/30. This resulted in 15 types of visual apparent motion. Based on the nearest-neighbor principle for the binding occurring during visual apparent motion (Shechter et al., 1988; Dawson, 1991), we expected participants to consistently perceive rightward visual apparent motion for an 8/30 shift while consistently perceiving leftward visual apparent motion for a 22/30 shift. A 15/30 shift was expected to equally often induce a left and rightward visual apparent motion percept (see **Figure 1**).

Trials contained one visual apparent motion stimulus. At the end of each trial, participants indicated in a two-alternative forced-choice regime which visual motion direction they had seen by pressing the left and right mouse button when they perceived left and rightward visual motion, respectively. Responses were made using a conventional PC mouse. During the response period, the gray disks disappeared and the next trial would start 500 ms after a response was made. Participants were instructed to fixate the green disk that persisted throughout the entire session at the center of the image (radius = 0.4°, luminance = 11.3 cd/$m^2$).

Visual apparent motion was presented in three different auditory conditions. During the no-sound condition, visual apparent motion was presented without auditory stimulation. During the left and rightward auditory motion conditions, visual apparent motion was presented in conjunction with auditory apparent motion stimuli. Leftward auditory motion trials contained white-noise bursts (82–85 dB) which were presented for 80 ms with an inter-stimulus interval of 20 ms at the far-right, mid-right, mid-left, and far-left speaker consecutively and vice versa for rightward auditory motion trials. Critically, the third sound was presented simultaneously with the onset of the second visual apparent motion component, which leads to an overlap of perceived visual and auditory motion in space and time.

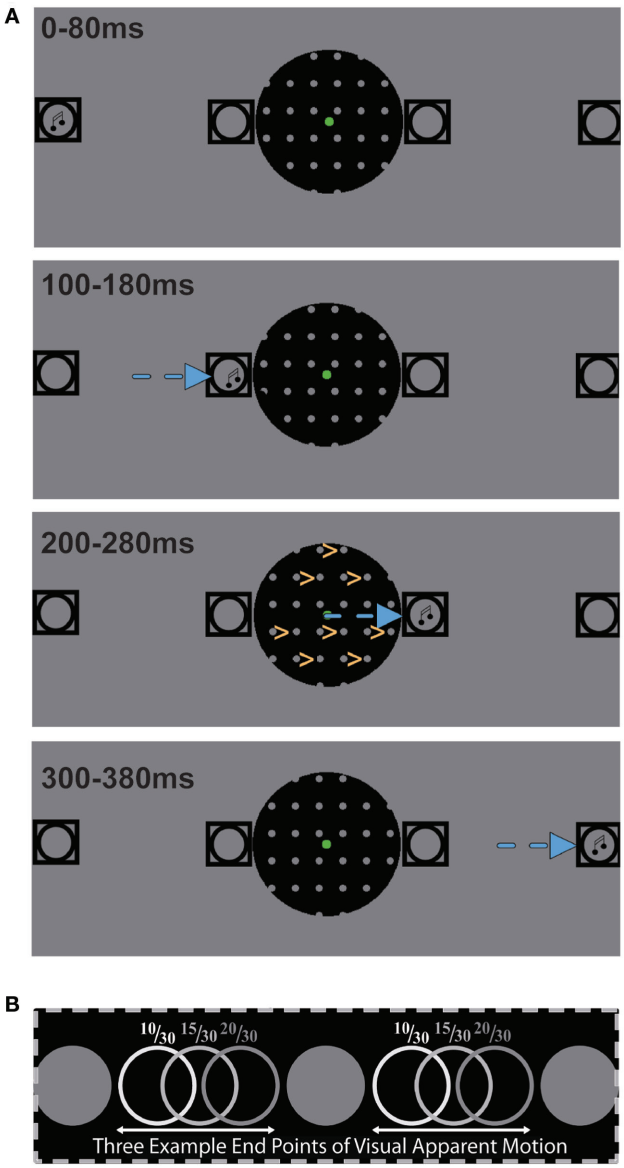

**FIGURE 1 | (A)** A schematic depiction of the stimuli employed in this study illustrating a rightward auditory motion trial during which the participant perceived rightward visual apparent motion. Squares depict the four speaker locations and the musical note symbols represent the location at which a sound is being presented. The direction of auditory apparent motion is depicted by the blue arrows and the direction of visual apparent motion by yellow arrows. **(B)** An illustration of three possible endpoints of visual apparent motion whose position is expressed as the distance from the nearest left starting point divided by the inter-column distance. The 10/30 endpoints most likely induce rightward visual apparent motion and the 20/30 endpoints most likely induce leftward visual apparent motion while the 15/30 endpoints are as likely to induce left as rightward visual apparent motion.

In total, each participant was presented 10 times with each of the 15 visual apparent motion stimuli for each trial type in each of two sessions (900 trials in total). The order of trials was fully randomized. On average each session took around 15 min and participants were allowed to take a short break in between the two sessions. Participants were told before the experiment that sounds would be presented during the experiment but that they were irrelevant for the visual motion task.

### *Control experiment*

Participants were seated in a dimly lighted room 55.5 cm away from a 19″ Samsung LCD monitor (width: 37.3°, height of 30.2°) that was part of an SMI Remote eye-tracking Device (RED 4.2-911-138). During the experiment they rested their heads on a chin and forehead rest to ensure a constant eye-to-screen distance. The eye tracker measured the location and radius of the left and right pupil at 50 Hz using two infrared light sources and an infrared camera. The horizontal and vertical gaze position was calculated based on the average left and right pupil position. We used SMI's eye tracker software iView X (v2.8) and calibrated the eye tracker using the standard RED nine-point calibration procedure.

The same type of visual stimulus was presented on this screen as during the main experiment. The size of the stimuli, however, differed. The radius of the aperture was 9.3°, the gray disks had a radius of 0.5°, the inter-column distance of the disk array was 3.0°, the inter-row distance of the disks was 3.5°, and the green fixation disk had a radius of 0.6°. During this experiment we used a 5.1 speaker set (Creative Inspire 16160); we placed four satellite speakers on a line 5 cm in front of the LCD monitor that ran parallel to the monitor. The speakers were set up at a height that put them on the same level as the horizontal midline of the LCD monitor. The far-left speaker was located 49.2° left from the center of the monitor and the mid-left speaker 16.4° left from the center of the monitor. The far-right speaker was located 49.2° right from the center of the monitor and the mid-right speaker 16.4° right from the center of the monitor. The inter-speaker distance was 32.8° between all adjacent speakers. We used the same noise bursts as during the main experiment (82–85 dB) and the timing of the sounds and visual stimuli was the same as for the main experiment. All visual stimuli and sounds were generated using Presentation software (version 14.1, Neurobehavioral Systems, Inc.).

Participants received the same task instructions as for the main experiment. We increased the number of stimuli per trial type to 30 (1350 trials in total). Subjects completed all trials during one session that took about 1 h.

## BEHAVIORAL ANALYSIS

Behavioral data from the main and the control experiment was analyzed in the same way. For each participant, we calculated the percentage of rightward visual apparent motion reports for each of the rightward shifts separately for the three main conditions. We fitted a logistic psychometric function to this data for each participant and condition separately (Draper and Smith, 1981). The tails of each logistic function were fixed to 0 and 100%, respectively. An exemplary participant's data and its fitted curves are shown in **Figure 2**. The fitted logistic function for the no-sound condition was used to estimate the shift that induced left and rightward visual apparent motion with an equal probability without auditory influences. This was realized by determining the shift at which the no-sound curve intersected with a 50% level of rightward visual apparent motion perception (V50 of the logistic function). This shift we refer to as the bistability shift and we determined this shift for each participant separately.

The main aim of this study was to test whether the direction of auditory motion affects the ratio of perceived left and rightward visual apparent motion when visual apparent motion is bistable. To this end, we estimated the proportion at which participants perceived rightward visual motion for the two main auditory conditions when a visual stimulus is presented using the individually defined bistability shift. This was realized by fitting two logistic functions to the data from left and rightward auditory motion trials for each participant separately and comparing the amplitude of these curves for the bistability shift (see **Figure 2**). If our hypothesis is correct, then these values should be higher than 50% for the rightward auditory motion condition and lower than 50% for the leftward auditory motion condition. We assessed the effect of auditory motion direction on the percentage reported rightward visual apparent motion during bistable visual stimulation in a repeated-measures analysis of variance (ANOVA) and subsequent paired $t$-tests. We furthermore assessed in another ANOVA the effect of auditory motion direction on the threshold and slope of the fitted logistic functions.

As an alternative analysis approach on the group level, we used a bootstrap technique implemented in the "Palamedes" toolbox for Matlab (Prins and Kingdom, 2009). Data of participants were combined for estimates of the psychometric functions for the three conditions leftward, no-sound, and rightward resulting in 380 trials (19 participants times 20 trials) per aspect-ratio value for the main experiment and 300 (10 participants times 30 trials) for the control experiment. The toolbox contains a function "PAL_PFLR_ModelComparison" that allows for comparisons between models with varying constraints on the model parameters for the different conditions. In our case, we wanted to know whether a model with differing thresholds or slopes would demonstrate a more adequate fit to the data than a model with fixed parameters across conditions. As in the other analyses, the logistic function was used as model function. We used the "Palamedes" toolbox to obtain maximum-likelihood estimates in order to identify the best-fitting parameters. Guess and lapse rates were fixed at 0 for the fitting procedures. For model comparison with the "PAL_PFLR_ModelComparison" function, likelihood ratios between the unconstrained and constrained model for the observed data are compared to a distribution of likelihood ratios obtained via Monte-Carlo simulations from the constrained model. We used 5000 simulations for our calculations and performed model comparisons for threshold and slope using all three conditions as well as pair-wise comparisons between conditions for the threshold parameter.

# RESULTS

## BEHAVIORAL RESULTS

### *Main experiment*

The average rightward shift at which the visual apparent motion stimulus was estimated to be bistable was 0.495 (SD = 0.0126) of

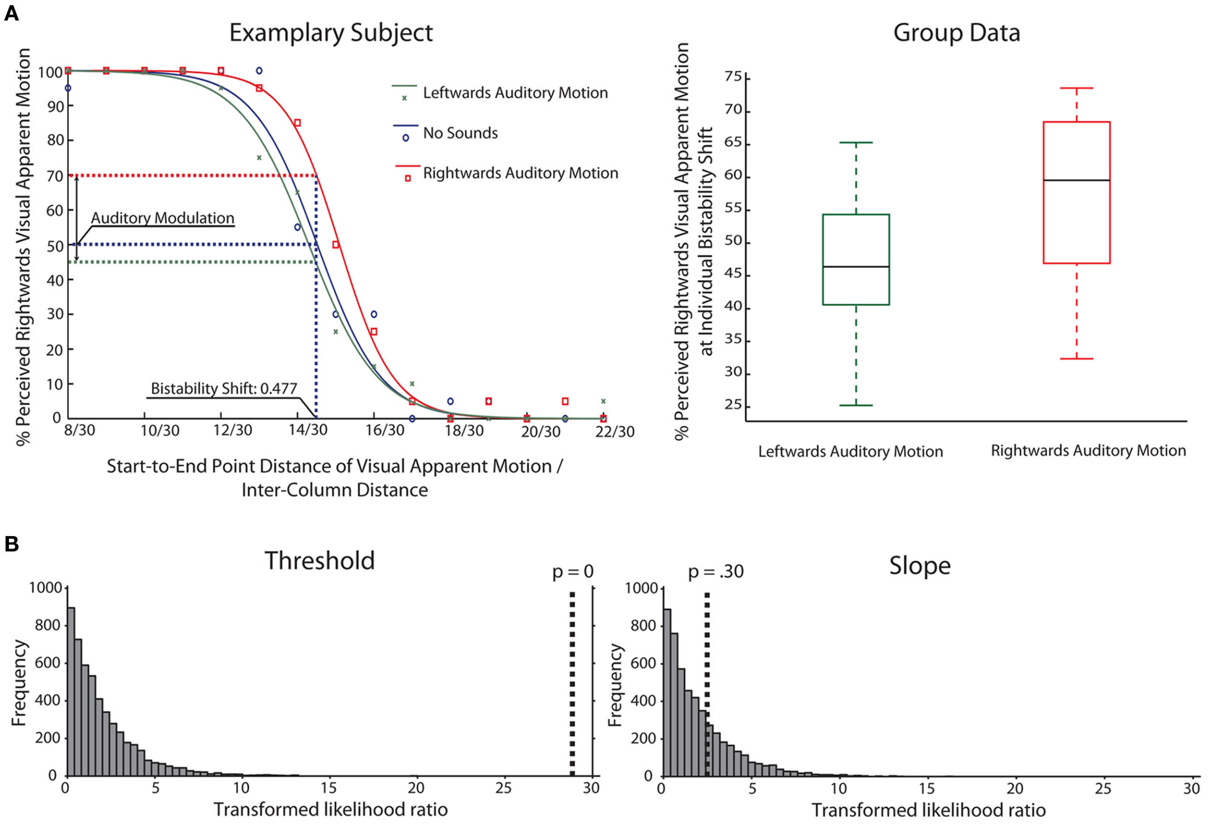

**FIGURE 2 | (A)** *Left*: This figure contains the plotted data for an exemplary participant (main experiment) and the fitted logistic functions for all three conditions. The visual apparent motion percept of this participant was estimated to be bistable for a shift of 0.477 of the inter-column distance. For this bistability shift the participant was estimated to perceive the visual apparent motion stimulus as moving rightward in 70% of the rightward auditory motion trials and in 46% of the leftward auditory motion trials. *Right*: Box-and-whisker diagrams depicting the group data for the main experiment for the percent perceived rightward visual apparent motion at the bistability shift which was estimated on an individual level. **(B)** Results of bootstrapping analysis ($N = 5000$) for group data comparing different models for thresholds and slopes in the three conditions (Leftward, No-Sound, Rightward). The histogram shows the distribution of likelihood ratios for the simulated data (see Materials and Methods for details). The dotted line indicates the likelihood ratio for the actual data. The results indicate that assuming varying thresholds for the three conditions is a more adequate model than a fixed threshold for all three conditions (*left panel*), whereas this is not the case for slope (*right panel*).

the inter-column distance. An ANOVA showed that the direction of auditory motion significantly affected the percentage of trials in which a bistable visual stimulus was seen as moving rightward ($F = 7.8, p < 0.003$ Greenhouse–Geisser). The bistable visual stimulus was estimated to be perceived as moving rightward in 57.2% (SD = 13.0%) of the rightward auditory motion trials while being perceived as moving rightward only in 47.0% (SD = 11.2%) of the leftward auditory motion trials. Paired $t$-tests showed that bistable visual apparent motion induced a rightward visual motion percept more often during rightward auditory motion trials as compared to leftward auditory motion trials ($t = 4.3$, $p < 0.001$ two-sided) and no-sound trials ($t = 2.4$, $p < 0.03$ two-sided). The difference between the leftward auditory motion condition and the no-sound condition, however, did not reach significance ($t = 1.2$, $p = 0.26$).

Auditory stimulation was found to have an effect on the threshold of the fitted logistic function ($F = 7.6$, $p < 0.004$ Greenhouse–Geisser) but not on the logistic function's slope ($F = 0.431$, $p = 0.65$ Greenhouse–Geisser). The V50 of the logistic function was 0.492 (SD = 0.0155), 0.495 (SD = 0.0126), and 0.505 (SD = 0.0146) of the inter-column distance for the leftward auditory motion condition, the no-sound condition, and the rightward auditory motion condition, respectively. The differences in threshold between the left and rightward auditory motion condition and between the rightward auditory motion condition and the no-sound condition reached significance (respectively: $t = 3.5$, $p < 0.004$ two-sided; $t = 2.5$, $p < 0.03$ two-sided) while the difference between the leftward auditory motion condition and the no-sound condition was not significant ($t = 1.2$, $p = 0.24$ two-sided). These findings indicate that the effects of auditory motion found in our main analysis are due to a shift of the sigmoidal psychometric function without a change of slope.

To confirm our results with an alternative analysis approach, we used a bootstrap technique implemented in the "Palamedes" toolbox (Prins and Kingdom, 2009). In short, the toolbox compares different models fitted to the psychophysical data and obtains significance values through Monte-Carlo simulations (see Materials and Methods for details and Kingdom and Prins, 2010). In our case, we compared models where either threshold or slope values were fixed across conditions (corresponding to the Null Hypothesis that all thresholds/slopes are equal across conditions) to models where parameters could vary freely between conditions. The models were compared on the group level, combining all participants' data for a pooled estimate.

There was a highly significant effect for the threshold parameter ($p < 0.001$; in fact, none of the simulated values was larger than the observed value, making $p$ effectively 0, see **Figure 2**), meaning that assuming the same threshold for the different conditions was an inadequate model of the observed data. Performing the model comparison for pairs of conditions, simulations showed that the threshold for the rightward condition was different from both the leftward and no-sound condition ($p < 0.001$). But no difference between leftward and no-sound conditions was found ($p = 0.30$) confirming the above mentioned analyses. In contrast, for the slope parameter, assuming the same slope values across conditions was an adequate model of the observed data ($p = 0.30$).

### *Control experiment*

The average rightward shift at which the visual apparent motion stimulus was estimated to be bistable was 0.497 (SD = 0.0105) of the inter-column distance. An ANOVA showed that the direction of auditory motion significantly affected the percentage of trials in which a bistable visual stimulus was seen as moving rightward ($F = 34.6$ $p < 0.001$ Greenhouse–Geisser). The bistable visual stimulus was estimated to be perceived as moving rightward in 68.5% (SD = 15.0%) of the rightward auditory motion trials while being perceived as moving rightward only in 42.8% (SD = 11.7%) of the leftward auditory motion trials. Paired $t$-tests showed that bistable visual apparent motion induces a rightward visual motion percept more often during rightward auditory motion trials as compared to leftward auditory motion trials ($t = 5.9$, $p < 0.001$ two-sided) and no-sound trials ($t = 3.9$, $p < 0.005$ two-sided). The difference between the leftward auditory motion condition and the no-sound condition, however, did not reach significance ($t = 1.9$, $p = 0.09$).

Auditory stimulation was found to have an effect on the threshold of the fitted logistic function ($F = 12.3$, $p < 0.002$ Greenhouse–Geisser) but not on the logistic function's slope ($F = 1.4$, $p = 0.27$ Greenhouse–Geisser). The V50 of the logistic function was 0.493 (SD = 0.0134), 0.497 (SD = 0.0105), and 0.510 (SD = 0.009) of the inter-column distance for the leftward auditory motion condition, the no-sound condition, and the rightward auditory motion condition, respectively. The differences in threshold between the left and rightward auditory motion condition and between the rightward auditory motion condition and the no-sound condition reached significance (respectively: $t = 4.1$, $p < 0.004$ two-sided; $t = 3.3$, $p < 0.02$ two-sided) while the difference between the leftward auditory motion condition and the no-sound condition was not significant ($t = 1.7$, $p = 0.13$ two-sided).

As for the main experiment, we also applied the bootstrapping approach to the data of the control experiment. In the model comparison, there was a highly significant effect of the threshold parameter ($p < 0.001$; again, none of the simulated values was larger than the observed value, making $p$ effectively 0, see **Figure 3**), meaning that assuming the same threshold for the different conditions was an inadequate model of the observed data. Performing the model comparison for pairs of conditions, simulations showed that the threshold for the rightward condition was different from both the leftward and no-sound condition ($p < 0.001$). The difference between leftward and no-sound conditions was marginally significant ($p = 0.05$). In the control experiment, there was also a significant effect of the slope parameter ($p = 0.003$). The slope was steeper for the rightward condition than the leftward ($p = 0.001$) or no-sound condition ($p = 0.007$); there was no slope difference between the leftward and no-sound condition ($p = 0.53$).

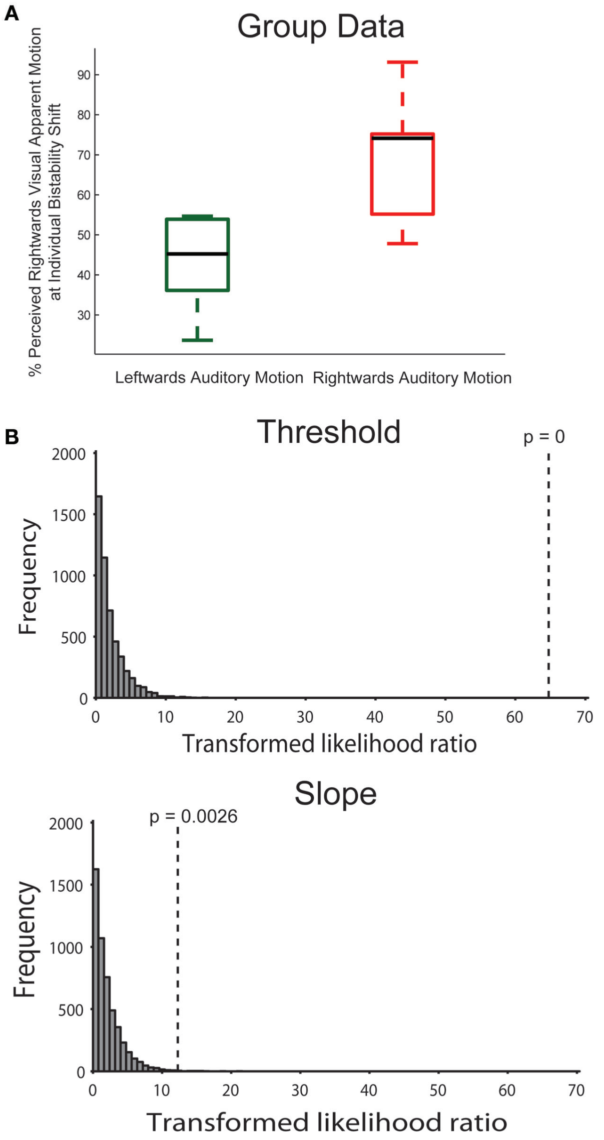

**FIGURE 3 | (A)** Same as **Figure 2A** right but showing the results for the control experiment. **(B)** In the control experiment, the bootstrapping analysis (see **Figure 2B**) confirmed the threshold effect. Additionally, this analysis identified a difference in the slope parameter of the psychometric function.

## EYE-TRACKING ANALYSIS AND RESULTS

We analyzed horizontal and vertical gaze angle for epochs from 400 ms before trial onset until 1300 ms after trial onset using custom-made Matlab code. Time points were labeled as blinks if the pupil diameter was smaller than one fourth of the median pupil diameter or when the vertical gaze position deviated more than six visual degrees from the fixation dot. These time points and the time within 60 ms from these time points were treated as missing values and excluded from subsequent analyses. To correct for imperfections of the calibration we redefined the horizontal and vertical gaze angle for the fixation dot for each participant as the mean of all of the participant's valid horizontal and vertical gaze angles.

For each participant and condition we calculated fixation performance by determining the percentage of time that his or her gaze angle deviated less than 1° visual angle from the fixation dot position. On average fixation performance was 71.0, 70.4, and 71.3% for the no-sound, leftward, and rightward auditory motion conditions respectively. Two-sided paired sample $t$-tests indicated that there was no significant effect of condition on fixation performance ($p > 0.05$ for all comparisons). We also created density plots for each condition after combining the data across all 10 participants (**Figure 4A**). These density plots also did not show any effect of condition on the distribution of gaze angle.

Additionally, we computed average event-related horizontal eye movements (ERHEMs) for each subject, sound condition, and visual apparent motion percept. First, we determined whether ERHEMs differ between trials during which participants were presented with left and rightward moving sounds. To determine this independently from the effect of visual apparent motion percept we averaged the ERHEMs for left and rightward visual apparent motion percept trials separately for the ERHEMs based on left and rightward moving sound trials. This resulted in an ERHEM for left and rightward sound motion for each subject (see **Figure 4B** for the grand mean and SEM). We tested for each 100 ms time bin after trial onset if there was a significant difference between the ERHEMs using a two-sided paired sample $t$-test ($p < 0.05$) across participants. We did not observe a significant difference between these two ERHEMs for any of the 100-ms time bins which suggests that sound motion direction had no effect on horizontal gaze angle.

Second, we assessed if ERHEMs differed between trials during which subjects perceived right and leftward visual apparent motion. To determine this independently from effects of sound condition we averaged the ERHEMs for the left and rightward sound conditions separately for the ERHEMs based on left and rightward visual apparent motion trials. This resulted in an ERHEM for left and rightward visual apparent motion percept for each subject (see **Figure 4B** for the grand mean and SEM). We observed a significant effect of visual apparent motion percept for the three time bins between 600 and 900 ms after trial onset. During this period, participant's gaze was on average shifted more to the right when participants perceived rightward visual apparent motion as compared to when they perceived leftward visual apparent motion. Note that this effect occurs 200 ms after all sounds and visual apparent motion stimuli were presented.

## DISCUSSION

The results of the main experiment show that the presence of rightward auditory motion causes visual stimuli to be perceived more often as moving rightward when the visual stimulus direction is ambiguous. Furthermore, we observed an increased likelihood of perceiving leftward visual motion when comparing trials containing leftward and rightward auditory motion. However, the ratio of left and rightward visual apparent motion perception was not found to differ between trials during which leftward auditory motion was presented and trials without auditory stimulation.

The findings of the main experiment were replicated by the control experiment during which we also measured eye movements. Again, visual motion perception was influenced in the direction of auditory motion. In addition to a threshold shift, there was also a hint toward a steeper slope of the psychometric function in the rightward condition. Concerning eye movements, we observed a shift in average horizontal gaze direction in the direction of the perceived visual apparent motion. This effect occurred only after all visual stimuli and sounds had been presented. After factoring out the effect of visual apparent motion percept, sound direction on its own was not found to have an effect on horizontal gaze direction. Therefore, it appears unlikely that the effect of sound motion on visual apparent motion perception observed in this study is mediated by sound-evoked eye movements.

Our findings are in contrast to claims made by previous studies that visual motion can capture auditory motion but not the other way around (Soto-Faraco et al., 2002, 2003, 2004, 2005; Sanabria et al., 2007). The lack of evidence in these studies for an effect of auditory motion on visual motion perception is most likely due to the fact that those studies employed visual stimuli that more reliably conveyed information about the direction of motion than the auditory stimuli. If audiovisual integration is based on Bayesian inference (Battaglia et al., 2003; Deneve and Pouget, 2004; Knill and Pouget, 2004), then bistable visual motion stimuli can be expected to be more susceptible to auditory influences than visual stimuli with unequivocal motion direction.

Our findings further support the emerging view that auditory motion stimuli can affect visual motion perception (Meyer and Wuerger, 2001; Meyer et al., 2005; Brooks et al., 2007; Freeman and Driver, 2008; Hidaka et al., 2009, 2011; Conrad et al., 2010; Teramoto et al., 2010). We propose that such an effect of moving sounds on visual motion perception results from sounds affecting neural activity in the visual cortex. This would be in line with the finding that auditory motion enhances visual responses in the visual motion complex hMT/V5+ when its direction is congruent with that of visual motion stimuli (Alink et al., 2008). Moreover, this result is supported by more recent studies showing that the direction of moving sounds can be decoded from activation patterns in the visual cortex of sighted (Alink et al., in press) and blind subjects (Wolbers et al., 2010) and that hMT/V5+ of congenitally blind subjects responds selectively to sound motion (Saenz et al., 2008; Bedny et al., 2010). Whether such effects of sound motion on the visual cortex actually contribute to the quality of auditory motion perception remains to be shown. One way of addressing this question would be assessing the effect of disruptive transcranial magnetic stimulation over hMT/V5+ on auditory motion perception.

To summarize, the current study shows that the direction of auditory motion can bias the direction of our visual motion

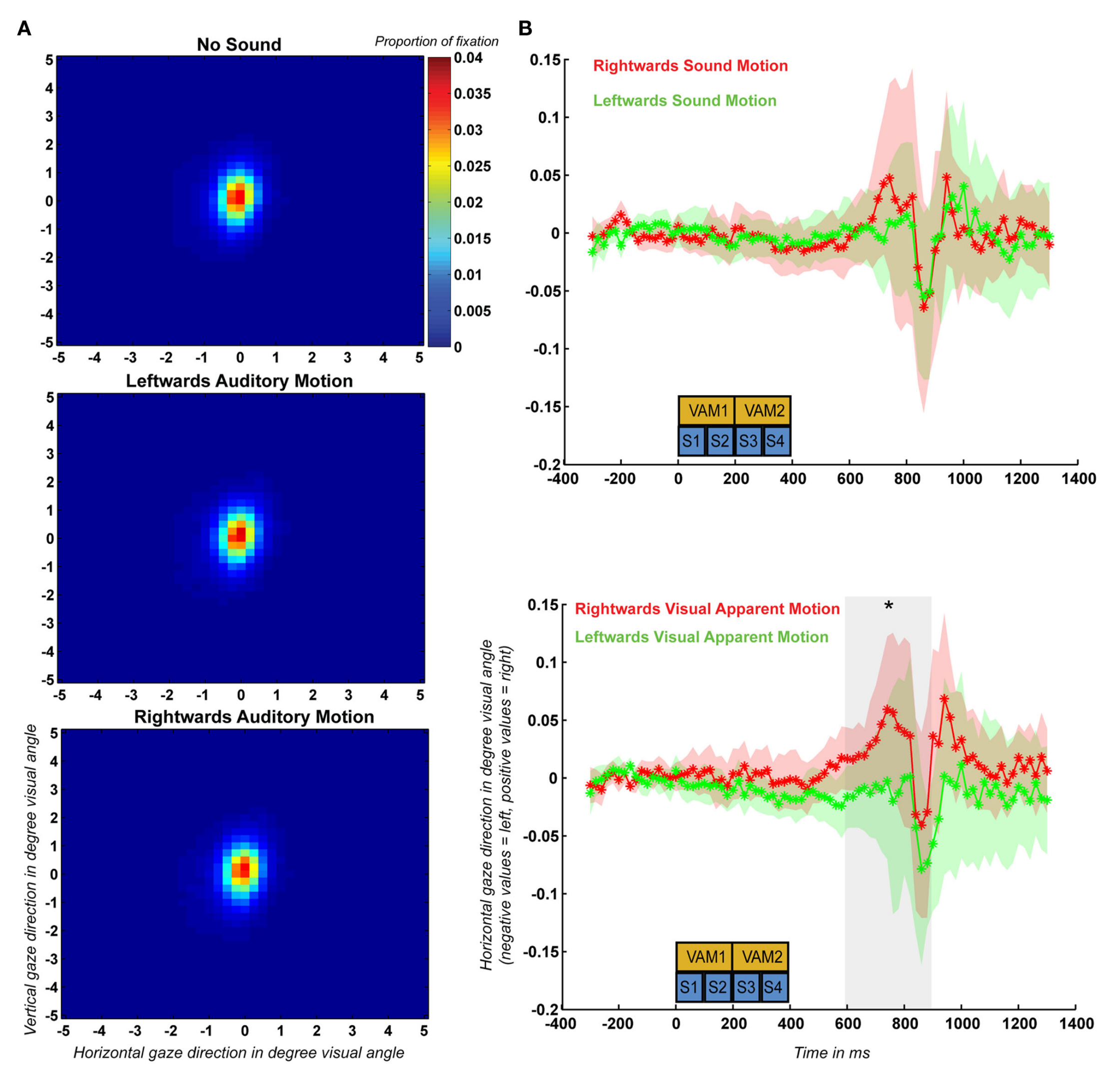

**FIGURE 4 | (A)** Spatial distribution of fixation recorded during the control experiment based on the eye-tracking data of all 10 participants plotted separately for the no-sound, leftward auditory motion, and rightward auditory motion conditions. **(B)** Mean and SEM (across participants) of the event-related horizontal eye movements (ERHEMs) computed for trials during which participants were presented with left and rightward auditory motion stimuli (*upper graph*) and mean and SEM of the ERHEMs computed for trials during which participants perceived left and rightward visual apparent motion (*lower graph*). The blue boxes (S1–S4) indicate the time at which the four sounds were presented during a trial and the yellow boxes (VAM1–VAM2) indicate the timing of the first and second visual apparent motion stimuli. The gray transparent box in the upper part indicates the time period (600–900 ms) during which there was a significant effect of visual apparent motion percept.

percept when visual motion direction is ambiguous. We did not observe an effect of sound motion direction on eye movements. Therefore, it appears that the effect of sound motion on visual motion perception takes place at a perceptual level rather than being induced by changes in eye position. Such a perceptual interaction is in line with several neurophysiological findings indicating that sound motion affects activation in the visual cortex.

## ACKNOWLEDGMENTS

Research funded by Netherlands Organisation for Scientific Research (NWO), Grant number: 825.10.023.

## REFERENCES

Alais, D., and Burr, D. (2004). No direction-specific bimodal facilitation for audiovisual motion detection. *Brain Res. Cogn. Brain Res.* 19, 185–194.

Alink, A., Euler, F., Kriegeskorte, N., Singer, W., and Kohler, A. (in press). Auditory motion direction encoding in auditory cortex and high-level visual cortex. *Hum. Brain Mapp.* doi: 10.1002/hbm.21263

Alink, A., Singer, W., and Muckli, L. (2008). Capture of auditory motion by vision is represented by an activation shift from auditory to visual motion cortex. *J. Neurosci.* 28, 2690.

Arrighi, R., Marini, F., and Burr, D. (2009). Meaningful auditory information enhances perception of visual biological motion. *J. Vis.* 9, 1–7.

Battaglia, P. W., Jacobs, R. A., and Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *J. Opt. Soc. Am. A* 20, 1391–1397.

Bedny, M., Konkle, T., Pelphrey, K., Saxe, R., and Pascual-Leone, A. (2010). Sensitive period for a multimodal response in human visual motion area MT/MST. *Curr. Biol.* 20, 1900–1906.

Brooks, A., Van Der Zwan, R., Billard, A., Petreska, B., Clarke, S., and Blanke, O. (2007). Auditory motion affects visual biological motion processing. *Neuropsychologia* 45, 523–530.

Carter, O., Konkle, T., Wang, Q., Hayward, V., and Moore, C. (2008). Tactile rivalry demonstrated with an ambiguous apparent-motion quartet. *Curr. Biol.* 18, 1050–1054.

Conrad, V., Bartels, A., Kleiner, M., and Noppeney, U. (2010). Audiovisual interactions in binocular rivalry. *J. Vis.* 10, 1–15.

Dawson, M. R. W. (1991). The how and why of what went where in apparent motion: modeling solutions to the motion correspondence problem. *Psychol. Rev.* 98, 569–603.

Deneve, S., and Pouget, A. (2004). Bayesian multisensory integration and cross-modal spatial links. *J. Physiol. Paris* 98, 249–258.

Draper, N. R., and Smith, H. (1981). *Applied Regression Analysis.* New York: Wiley.

Freeman, E., and Driver, J. (2008). Direction of visual apparent motion driven solely by timing of a static sound. *Curr. Biol.* 18, 1262–1266.

Hidaka, S., Manaka, Y., Teramoto, W., Sugita, Y., Miyauchi, R., Gyoba, J., Suzuki, Y., and Iwaya, Y. (2009). Alternation of sound location induces visual motion perception of a static object. *PLoS ONE* 4, e8188. doi:10.1371/journal.pone.0008188

Hidaka, S., Teramoto, W., Sugita, Y., Manaka, Y., Sakamoto, S., Suzuki, Y., and Coleman, M. (2011). Auditory motion information drives visual motion perception. *PLoS ONE* 6, e17499. doi:10.1371/journal.pone.0017499

Kingdom, F. A. A., and Prins, N. (2010). *Psychophysics: A Practical Introduction.* London: Academic Press.

Kitagawa, N., and Ichihara, S. (2002). Hearing visual motion in depth. *Nature* 416, 172–174.

Knill, D. C., and Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* 27, 712–719.

Laubrock, J., Engbert, R., and Kliegl, R. (2005). Microsaccade dynamics during covert attention. *Vision Res.* 45, 721–730.

Meyer, G. F., and Wuerger, S. M. (2001). Cross-modal integration of auditory and visual motion signals. *Neuroreport* 12, 2557.

Meyer, G. F., and Wuerger, S. M., Röhrbein, F., and Zetzsche, C. (2005). Low-level integration of auditory and visual motion signals requires spatial co-localisation. *Exp. Brain Res.* 166, 538–547.

Prins, N., and Kingdom, F. A. A. (2009). *Palamedes: Matlab Routines for Analyzing Psychophysical Data.* Available at: http://www.palamedestoolbox.org

Saenz, M., Lewis, L. B., Huth, A. G., Fine, I., and Koch, C. (2008). Visual motion area MT+/V5 responds to auditory motion in human sight-recovery subjects. *J. Neurosci.* 28, 5141.

Sanabria, D., Spence, C., and Soto-Faraco, S. (2007). Perceptual and decisional contributions to audiovisual interactions in the perception of apparent motion: a signal detection study. *Cognition* 102, 299–310.

Shechter, S., Hochstein, S., and Hillman, P. (1988). Shape similarity and distance disparity as apparent motion correspondence cues. *Vision Res.* 28, 1013–1021.

Soto-Faraco, S., Kingstone, A., and Spence, C. (2003). Multisensory contributions to the perception of motion. *Neuropsychologia* 41, 1847–1862.

Soto-Faraco, S., Lyons, J., Gazzaniga, M., Spence, C., and Kingstone, A. (2002). The ventriloquist in motion: illusory capture of dynamic information across sensory modalities. *Brain Res. Cogn. Brain Res.* 14, 139–146.

Soto-Faraco, S., Spence, C., and Kingstone, A. (2004). Cross-modal dynamic capture: congruency effects in the perception of motion across sensory modalities. *J. Exp. Psychol. Hum. Percept. Perform.* 30, 330–345.

Soto-Faraco, S., Spence, C., and Kingstone, A. (2005). Assessing automaticity in the audiovisual integration of motion. *Acta Psychol. (Amst.)* 118, 71–92.

Teramoto, W., Manaka, Y., Hidaka, S., Sugita, Y., Miyauchi, R., Sakamoto, S., Gyoba, J., Iwaya, Y., and Suzuki, Y. (2010). Visual motion perception induced by sounds in vertical plane. *Neurosci. Lett.* 479, 221–225.

Williams, Z. M., Elfar, J. C., Eskandar, E. N., Toth, L. J., and Assad, J. A. (2003). Parietal activity and the perceived direction of ambiguous apparent motion. *Nat. Neurosci.* 6, 616–623.

Wolbers, T., Zahorik, P., and Giudice, N. A. (2010). Decoding the direction of auditory motion in blind humans. *Neuroimage* 56, 681–687.

Wuerger, S. M., Hofbauer, M., and Meyer, G. F. (2003). The integration of auditory and visual motion signals at threshold. *Atten. Percept. Psychophys.* 65, 1188–1196.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 14 June 2011; accepted: 12 December 2011; published online: 02 January 2012.*

*Citation: Alink A, Euler F, Galeano E, Krugliak A, Singer W and Kohler A (2012) Auditory motion capturing ambiguous visual motion. Front. Psychology* ***2****:391. doi: 10.3389/fpsyg.2011.00391*

*This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.*

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 05 April 2011
doi: 10.3389/fpsyg.2011.00056

# Vision and audition do not share attentional resources in sustained tasks

**Roberto Arrighi[1,2,3]*, Roy Lunardi[1] and David Burr[1,2,3]**

[1] Facoltà di Psicologia, Università degli Studi di Firenze, Firenze, Italy
[2] Istituto Scientifico per la Neuropsichiatria dell'Infanzia e dell'Adolescenza Fondazione Stella Maris, Pisa, Italy
[3] Neuroscience Institute, National Research Council, Pisa, Italy

**Edited by:**
*Nadia Bolognini, University of Milano-Bicocca, Italy*
*Angelo Maravita, University of Milano-Bicocca, Italy*

**Reviewed by:**
*Salvador Soto-Faraco, Pompeu Fabra University, Spain*
*Silvia Savazzi, University of Verona, Italy*

***Correspondence:**
*Roberto Arrighi, Department of Psychology, University of Florence, Via di san Salvi 12, Pad. 26, 50135 – Florence, Italy.*
*e-mail: roberto.arrighi@gmail.com*

Our perceptual capacities are limited by attentional resources. One important question is whether these resources are allocated separately to each sense or shared between them. We addressed this issue by asking subjects to perform a double task, either in the same modality or in different modalities (vision and audition). The primary task was a multiple object-tracking task (Pylyshyn and Storm, 1988), in which observers were required to track between 2 and 5 dots for 4 s. Concurrently, they were required to identify either which out of three gratings spaced over the interval differed in contrast or, in the auditory version of the same task, which tone differed in frequency relative to the two reference tones. The results show that while the concurrent visual contrast discrimination reduced tracking ability by about 0.7 d′, the concurrent auditory task had virtually no effect. This confirms previous reports that vision and audition use separate attentional resources, consistent with fMRI findings of attentional effects as early as V1 and A1. The results have clear implications for effective design of instrumentation and forms of audio–visual communication devices.

**Keywords: sustained attention, audio–visual integration, cross-modal perception**

## INTRODUCTION

To successfully interact with the stimuli of our environment, we need to process selectively the information most relevant for our tasks. This process is usually termed "attention" (James, 1890/1950). When stimuli are attended to their processing become more rapid, more accurate, and more detailed (Posner et al., 1980; Desimone and Duncan, 1995; Carrasco and McElree, 2001; Carrasco et al., 2004; Liu et al., 2005, 2009). Attention improves performance on several visual tasks, such as contrast sensitivity, speed and orientation discrimination as well as spatial resolution (Lee et al., 1999; Morrone et al., 2002; Carrasco et al., 2004; Alais et al., 2006a). As attentive resources are limited, when the stimuli demanding attention for a perceptual task exceed system capacity, performance decreases. For example, in a visual search task in which an object (target) has to be detected amongst irrelevant items (distractors), reaction times increase directly with distractor number (unless the difference between the stimuli is so striking to make the target pop out from the cluttered scene). This correlation reflects the limited capacity of selective attention that prevents the observer from monitoring all items at the same time.

Similarly, when more than one perceptual task is performed at the same time, overall performance decreases because of the underlying processing limitations. This occurs even for simple tasks, such as naming a word or identifying the pitch of a tone (Pashler, 1992; Pashler and O'Brien, 1993; Huang et al., 2004). Interference between concurrent perceptual tasks of the same sensory modality has been consistently reported in many psychological and psychophysical studies (Navon et al., 1984; Pashler, 1994; Bonnel and Prinzmetal, 1998; Alais et al., 2006b). However, the evidence for audiovisual cross-modal interference is conflicting (Duncan et al., 1997; Jolicoeur, 1999; Arnell and Duncan, 2002). Bonnel and Hafter (1998) found that in a identification task in which the sign of a change (luminance in vision and intensity in audition) had to be detected, performance in dual-task conditions were lower than in the single-task conditions regardless the interference was in the same or different modalities. Spence et al. (2000) found that selecting an auditory stream of words presented concurrently with a second (distractor) stream, it is more difficult if a video of moving lips mimicking the distracting sounds it is also displayed. These psychophysical findings are not only congruent with some of the cognitive literature of the 1970s and 1980s (Taylor et al., 1967; Tulving and Lindsay, 1967; Alais et al., 2006b), but also with recent neurophysiological and imaging results. For example, Joassin et al. (2004) examined the electrophysiological correlates for auditory interference with vision by an identification task of non-ambiguous complex stimuli such as faces and voices. Their results suggest that cross-modal interactions occur at various different stages, involving brain areas such as fusiform gyrus, associative auditory areas (BA 22), and the superior frontal gyri. Hein et al. (2007) showed with a functional magnetic resonance (fMRI) study, that even without competing motor responses, a simple auditory decision interferes with visual processing at neural levels including prefrontal cortex, middle temporal cortex, and other visual regions. Taken together these results imply that limitations on resources for vision and audition operate at a central level of processing, rather than in the auditory and visual peripheral senses.

However, much evidence also supports the notion of independence of attentional resources for vision and audition (Allport et al., 1972; Triesman and Davies, 1973; Shiffrin and Grantham, 1974; Alais et al., 2006b; Santangelo et al., 2010). For example, Larsen et al. (2003) compared subjects' accuracy for identification of two concurrent stimuli (such as a visual and spoken letter) relative to performance in a single-task. They found that the proportion of correct response was almost the same for all experimental conditions and, furthermore, in the divided-attention condition the probability to correctly report a stimulus in one modality was independent of whether the stimulus was correctly reported in the other modality. Similarly, Bonnel and Hafter (1998) used an audiovisual dual-task paradigm to show that when identification of the direction of a stimulus change is capacity-limited (see above), simple detection of visual and auditory patterns is governed by "capacity-free" processes, as in the detection task there was no performance drop compared with single-task controls. Similar results have been achieved by Alais et al. (2006b) by measuring discrimination thresholds for visual contrast and auditory pitch. Visual thresholds were unaffected by concurrent pitch discrimination of chords and vice versa. However, when two tasks were performed within the same modality, thresholds increased by a factor of around two for visual discrimination and four for auditory discrimination. In line with these psychophysical results, a variety of imaging studies suggests that attention can act unimodally at early levels including the primary cortices such as A1 and V1 (Jancke et al., 1999a,b; Posner and Gilbert, 1999; Somers et al., 1999).

Most of the studies mentioned deal with dual-task conditions where both tasks are brief (hundreds of milliseconds) stimuli to be detected or discriminated. Very few consider conditions in which one of the tasks must be performed by continuously monitoring a specific pattern over a temporal scale of seconds, even though this is a typical requirement for many everyday activities, such as reading or driving. These tasks require *sustained* rather than *transient* attention. Here we investigate whether sustained attentional resources are independent for vision and audition. We measure performance on the multiple object-tracking (MOT) task of Pylyshyn and Storm (1988), while asking subjects to perform simultaneously either a visual contrast discrimination task or an auditory pitch discrimination task. The results show strong within-modality interference, but very little cross-modality interference, strongly supporting the idea that in sustained tasks each modality has access to a separate pool of attentional resources.

## MATERIALS AND METHODS

### SUBJECTS

Four naive subjects (two males and two females, mean age 26 years), all with normal hearing and normal or corrected-to-normal visual acuity, served as subjects. All gave informed consent to participate to the study that was conducted in accordance with the guidelines of the University of Florence. The tasks were performed in a dimly lit, sound-attenuated room.

### STIMULI AND PROCEDURE

All visual stimuli were presented on a Sony Trinitron CRT monitor (screen resolution of 1024 × 768 pixels, 32 bit color depth, refresh rate of 60 Hz, and mean luminance 68.5 cd/m$^2$) subtending (40° × 30°) at the subjects view distance of 57 cm. To create visual stimuli we used Psychophysics toolbox (version 2) for MATLAB (Brainard, 1997; Pelli, 1997) on a Mac G4 running Mac OSX 9. Auditory stimuli were digitized at a rate of 65 kHz, and presented through two high quality loudspeakers (Creative MMS 30) flanking the computer screen and lying in the same plane 60 cm from the subject. Speaker separation was around 80 cm and stimuli intensity was 75 dB at the sound source.

Subjects were tested on two different kinds of perceptual tasks. The primary task was visual tracking of multiple moving objects (MOT; Pylyshyn and Storm, 1988). The MOT task consisted of 12 disks (diameter 0.9°) moving across a gray background at 5°/s. They moved in straight lines, and when colliding with other dots or the sides bounced appropriately (obeying the laws of physics). At the start of each trial 3–5 disks were displayed in green (*xyY* coordinates = 0.25,0.69, 39.5) for 2 s to indicate that those were the targets whilst the remaining were displayed an isoluminant red (*xyY* coordinates = 0.61, 0.33, 39.5). The trial continued for 4 s (tracking period), then disks stopped and four became orange (*xyY* coordinates = 0.52,0.44, 39.4; see **Movie S1** in Supplementary Material). The subjects' task was to choose which of these was the target (only one valid target turned orange on each trial). Subjects were familiarized with the task during a training session of 50 trials before starting the experimental protocol. Each experimental session had five trials per condition (varying in the number of dots to track) for a total of 15 trials per session. All subjects were tested for five sessions for a total of 75 trials. No feedback was provided, but subjects could check their overall performance at the end of each session.

Stimuli for the secondary visual task were luminance-modulated gratings of 0.5 s duration with a spatial frequency of 3 c/deg covering the entire screen. On each trial (4 s duration, during the dot tracking) subjects were presented with a sequence of three gratings, ramped in and out within a raised cosine envelope (over 20 ms), with an inter stimulus interval randomly chosen between 0.5 and 1.3 s. Two out of three gratings had the same contrast (50%) while the target grating (that subjects had to detect), randomly first, second or third in the sequence, had more or less contrast. The size of the contrast difference (Δ) was chosen from trial to trial by means of an adaptive staircase QUEST (Watson and Pelli, 1983) that homed in on threshold (67% of correct responses). The auditory secondary stimulus was of a sequence of three tones with the same presentation duration and temporal spacing as the visual version, two reference stimuli of 880 Hz with the target frequency differing from trial to trial by ±Δ Hz. In the dual-task condition subjects performed both the contrast or frequency discrimination task, and the MOT task. To avoid possible biases for response order we counterbalanced subjects responding first to the MOT task with those that responded for to the secondary task.

## RESULTS

To evaluate the costs of dividing attention between sensory modalities, we measured subject performance for visual tracking alone, or with either an auditory or a visual secondary task. **Figure 1** shows the individual results for the three experimental conditions, plotting performance (d′) in the dual-task conditions against single-task performance. Each small symbol indicates individual subject performance in a given condition defined by the number of dots to track whilst large circles indicate the data averaged across subjects and conditions. It is quite clear that the concurrent visual task greatly reduced performance, shown by the average decrease in d′ from 2.48 to 1.50, and also by the fact that all individual data

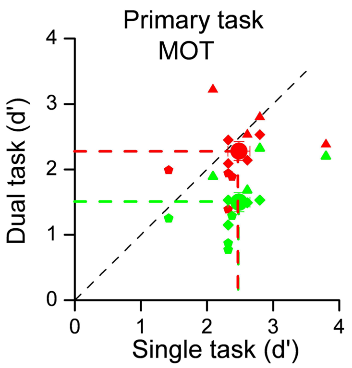

**FIGURE 1 | Sensitivities for the MOT task performed alone (on the abscissa) plotted against sensitivities for dual-task conditions (on the ordinate).** The 12 data points represent 4 subjects in 3 experimental conditions, defined by the number of dots to track (from 3 to 5). Green symbols refer to the intra-modal condition (secondary task contrast discrimination), red to the cross-modal condition (auditory secondary task). Small symbols refer to individual data (different symbols shape indicates different number of dots to track: three dots → triangles, four dots → diamond, and five dots → pentagon) whilst large symbols to averages. There is a clear effect for intra-modal interference, but not for cross-modal interference.

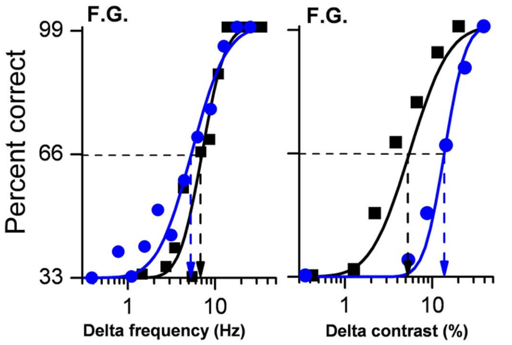

**FIGURE 2 | Psychometric functions for auditory frequency discrimination (left panel) and visual contrast discrimination (right panel) for subjects F.G.** Performance in the auditory task was almost identical when frequency discrimination was performed alone (black data points and lines) or together with a visual MOT task (blue data points and lines) as shown by the almost overlapping curves. However, when the two concurrent tasks were of the same sensory modality (vision), subject performance was dramatically reduced by around a factor of 3.

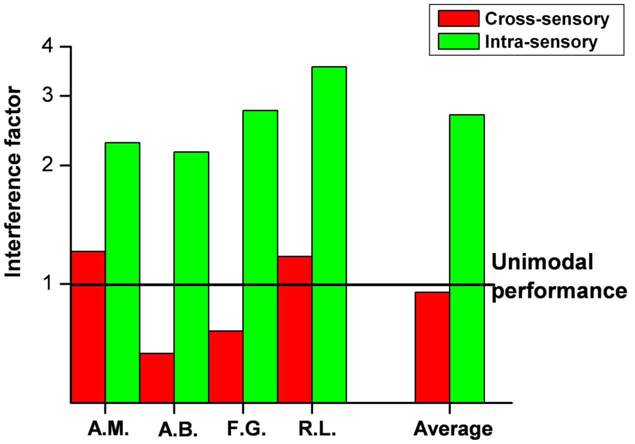

**FIGURE 3 | Subject performance on the secondary task, either auditory (red bars) or visual (green bars).** The interference factor is defined as the ratio between dual-task and single-task thresholds (a value of one meaning no interference between modalities.

lie below the equality line. The difference was highly significant (one-tailed paired $t$-test: $t_{11} = 6.98$, $p < 0.001$). However, when the competing task was auditory rather than visual, there was no effect on tracking performance. Average d′ was virtually unchanged (2.48 vs 2.28), certainly not significant (one-tailed paired $t$-test; $t_{11} = 1.07$, $p = 0.30$).

We also measured sensitivity for both the visual and auditory secondary tasks when performed alone and matched these results with those achieved in the dual-task condition. Examples of psychometric functions for subject F.G, are shown in **Figure 2**.

Auditory frequency discrimination is shown on the left, visual contrast discrimination on the right. It is obvious that the auditory discrimination was little affected by the concurrent visual tracking task. The two psychometric functions (best fitting cumulative Gaussian functions) are virtually identical, yielding thresholds (Δ frequency yielding 66% correct target identification) close to 6–7 Hz in both conditions. However, visual contrast discrimination thresholds were much higher in the dual than in the single-task condition, 5.1 compared with 14.5 (a factor of nearly three).

**Figure 3** plots for all subjects the interference factor (ratio of dual- to single-task thresholds) for the within and between modality conditions. It is clear that the auditory task is relatively immune to interference (average factor –0.05), while performance for luminance contrast discrimination thresholds increased by a factor of more than 2.5.

## DISCUSSION

In this paper we asked whether vision and audition share cognitive attentional resources in performing sustained tasks, particularly relevant for everyday functioning. As most previous research has been restricted to tasks spanning only a few milliseconds (Larsen et al., 2003; Alais et al., 2006b), or conditions with fast streams of simple auditory or visual patterns (Duncan et al., 1997), our study provides new knowledge about attentional mechanisms in ecological situations, where prolonged monitoring of information is necessary. The results clearly indicate that under these conditions, vision and audition have access to separate cognitive resources. Performance on a sustained task, typical of everyday requirement, was completely unaffected by a concurrent auditory discrimination task. The lack of interference did not reflect a bias in deploying attention on the visual primary task more than on the auditory task, as both tasks were performed as well as when they were presented alone. On the other hand, sharing attention between two tasks of the same sensory modality produces a robust decrease of performance for both primary and secondary tasks.

That vision and audition have access to separate cognitive resources is consistent with imaging studies showing that attention can modulate responses in primary and secondary visual and auditory cortexes (Gandhi et al., 1999; Jancke et al., 1999a,b; Somers et al., 1999). If in both modalities attentional effects modulate neural responses at these early of sensory information processing, when the visual and auditory signals are relatively independent, it is reasonable that few interactions are seen between these two senses.

Our results are important not only for the psychophysical data on the role of sustained attention between modalities, but also because they establish guidelines in designing audio–visual instrumentation. Information should be divided as much as possible between modalities, to maximize on the attentional resources available. This becomes increasingly more important as more virtual-reality applications are developed and are used routinely in everyday life.

## ACKNOWLEDGMENT

This study has been supported by Italian Ministry of Universities and Research and EC project "STANIB" (FP7 ERC).

## SUPPLEMENTARY MATERIAL

The Movie S1 for this article can be found online at http://www.frontiersin.org/Perception_Science/10.3389/fpsyg.2011.00056/abstract

## REFERENCES

Alais, D., Lorenceau, J., Arrighi, R., and Cass, J. (2006a). Contour interactions between pairs of Gabors engaged in binocular rivalry reveal a map of the association field. *Vision Res.* 46, 1473–1487.

Alais, D., Morrone, C., and Burr, D. (2006b). Separate attentional resources for vision and audition. *Proc. Biol. Sci.* 273, 1339–1345.

Allport, D. A., Antonis, B., and Reynolds, P. (1972). On the division of attention: a disproof of the single channel hypothesis. *Q. J. Exp. Psychol.* 24, 225–235.

Arnell, K. M., and Duncan, J. (2002). Separate and shared sources of dual-task cost in stimulus identification and response selection. *Cogn. Psychol.* 44, 105–147.

Bonnel, A. M., and Hafter, E. R. (1998). Divided attention between simultaneous auditory and visual signals. *Percept. Psychophys.* 60, 179–190.

Bonnel, A. M., and Prinzmetal, W. (1998). Dividing attention between the color and the shape of objects. *Percept. Psychophys.* 60, 113–124.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spat. Vis.* 10, 433–436.

Carrasco, M., Ling, S., and Read, S. (2004). Attention alters appearance. *Nat. Neurosci.* 7, 308–313.

Carrasco, M., and McElree, B. (2001). Covert attention accelerates the rate of visual information processing. *Proc. Natl. Acad. Sci. U.S.A.* 98, 5363–5367.

Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222.

Duncan, J., Martens, S., and Ward, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature* 387, 808–810.

Gandhi, S. P., Heeger, D. J., and Boynton, G. M. (1999). Spatial attention affects brain activity in human primary visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 96, 3314–3319.

Hein, G., Alink, A., Kleinschmidt, A., and Muller, N. G. (2007). Competing neural responses for auditory and visual decisions. *PLoS ONE* 2, e320. doi: 10.1371/journal.pone.0000320

Huang, L., Holcombe, A. O., and Pashler, H. (2004). Repetition priming in visual search: episodic retrieval, not feature priming. *Mem. Cognit.* 32, 12–20.

James, W. (1890/1950). *The Principles of Psychology*, Vol. 1. New York: Dover.

Jancke, L., Mirzazade, S., and Shah, N. J. (1999a). Attention modulates activity in the primary and the secondary auditory cortex: a functional magnetic resonance imaging study in human subjects. *Neurosci. Lett.* 266, 125–128.

Jancke, L., Mirzazade, S., and Shah, N. J. (1999b). Attention modulates the blood oxygen level dependent response in the primary visual cortex measured with functional magnetic resonance imaging. *Naturwissenschaften* 86, 79–81.

Joassin, F., Maurage, P., Bruyer, R., Crommelinck, M., and Campanella, S. (2004). When audition alters vision: an event-related potential study of the cross-modal interactions between faces and voices. *Neurosci. Lett.* 369, 132–137.

Jolicoeur, P. (1999). Restricted attentional capacity between sensory modalities. *Psychon. Bull. Rev.* 6, 87–92.

Larsen, A., McIlhagga, W., Baert, J., and Bundesen, C. (2003). Seeing or hearing? Perceptual independence, modality confusions, and crossmodal congruity effects with focused and divided attention. *Percept. Psychophys.* 65, 568–574.

Lee, D. K., Itti, L., Koch, C., and Braun, J. (1999). Attention activates winner-take-all competition among visual filters. *Nat. Neurosci.* 2, 375–381.

Liu, T., Abrams, J., and Carrasco, M. (2009). Voluntary attention enhances contrast appearance. *Psychol. Sci.* 20, 354–362.

Liu, T., Pestilli, F., and Carrasco, M. (2005). Transient attention enhances perceptual performance and FMRI response in human visual cortex. *Neuron* 45, 469–477.

Morrone, M. C., Denti, V., and Spinelli, D. (2002). Color and luminance contrasts attract independent attention. *Curr. Biol.* 12, 1134–1137.

Navon, D., Gopher, D., Chillag, N., and Spitz, G. (1984). On separability of and interference between tracking dimensions in dual-axis tracking. *J. Mot. Behav.* 16, 364–391.

Pashler, H. (1992). Attentional limitations in doing two tasks at the same time. *Curr. Dir. Psychol. Sci.* 1, 44–48.

Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. *Psychol. Bull.* 116, 220–244.

Pashler, H., and O'Brien, S. (1993). Dual-task interference and the cerebral hemispheres. *J. Exp. Psychol. Hum. Percept. Perform.* 19, 315–330.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437–442.

Posner, M. I., and Gilbert, C. D. (1999). Attention and primary visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 96, 2585–2587.

Posner, M. I., Snyder, C. R., and Davidson, B. J. (1980). Attention and the detection of signals. *J. Exp. Psychol.* 109, 160–174.

Pylyshyn, Z. W., and Storm, R. W. (1988). Tracking multiple independent targets: evidence for a parallel tracking mechanism. *Spat. Vis.* 3, 179–197.

Santangelo, V., Fagioli, S., and Macaluso, E. (2010). The costs of monitoring simultaneously two sensory modalities decrease when dividing attention in space. *Neuroimage* 49, 2717–2727.

Shiffrin, R. M., and Grantham, D. W. (1974). Can attention be allocated to sensory modalities? *Percept. Psychophys.* 15, 460–474.

Somers, D. C., Dale, A. M., Seiffert, A. E., and Tootell, R. B. (1999). Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 96, 1663–1668.

Spence, C., Ranson, J., and Driver, J. (2000). Cross-modal selective attention: on the difficulty of ignoring sounds at the locus of visual attention. *Percept. Psychophys.* 62, 410–424.

Taylor, M. M., Lindsay, P. H., and Forbes, S. M. (1967). Quantification of shared capacity processing in auditory and visual discrimination. *Acta Psychol. (Amst)* 27, 223–229.

Triesman, A. M., and Davies, A. (1973). *Divided Attention to Ear and Eye*. New York, NY: Academic Press.

Tulving, E., and Lindsay, P. H. (1967). Identification of simultaneously presented simple visual and auditory stimuli. *Acta Psychol. (Amst)* 27, 101–109.

Watson, A. B., and Pelli, D. G. (1983). QUEST: a Bayesian adaptive psychometric method. *Percept. Psychophys.* 33, 113–120.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 24 January 2011; accepted: 23 March 2011; published online: 05 April 2011.*

*Citation: Arrighi R, Lunardi R and Burr D (2011) Vision and audition do not share attentional resources in sustained tasks. Front. Psychology* ***2****:56. doi: 10.3389/fpsyg.2011.00056*

*This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.*

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 12 September 2011
doi: 10.3389/fpsyg.2011.00212

# Crossmodal constraints on human perceptual awareness: auditory semantic modulation of binocular rivalry

***Yi-Chuan Chen*[1]*, Su-Ling Yeh*[2] *and Charles Spence*[1]***

[1] *Department of Experimental Psychology, University of Oxford, Oxford, UK*
[2] *Department of Psychology, National Taiwan University, Taipei, Taiwan*

***Edited by:***
*Nadia Bolognini, University Milano-Bicocca, Italy*

***Reviewed by:***
*Toemme Noesselt, Otto von Guericke University, Germany*
*Davide Bottari, University of Hamburg, Germany*

****Correspondence:***
*Yi-Chuan Chen, Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK.*
*e-mail: yi-chuan.chen@queens.oxon.org*

We report a series of experiments utilizing the binocular rivalry paradigm designed to investigate whether auditory semantic context modulates visual awareness. Binocular rivalry refers to the phenomenon whereby when two different figures are presented to each eye, observers perceive each figure as being dominant in alternation over time. The results demonstrate that participants report a particular percept as being dominant for *less* of the time when listening to an auditory soundtrack that happens to be semantically congruent with the other alternative (i.e., the *competing*) percept, as compared to when listening to an auditory soundtrack that was irrelevant to both visual figures (Experiment 1A). When a visually presented word was provided as a semantic cue, no such semantic modulatory effect was observed (Experiment 1B). We also demonstrate that the crossmodal semantic modulation of binocular rivalry was robustly observed irrespective of participants' attentional control over the dichoptic figures and the relative luminance contrast between the figures (Experiments 2A and 2B). The pattern of crossmodal semantic effects reported here cannot simply be attributed to the meaning of the soundtrack guiding participants' attention or biasing their behavioral responses. Hence, these results support the claim that crossmodal perceptual information can serve as a constraint on human visual awareness in terms of their semantic congruency.

**Keywords: multisensory, audiovisual interaction, semantic congruency, consciousness, attention, stimulus contrast**

## INTRODUCTION

When viewing a scene, visual background context provides useful semantic information that can improve the identification of a visual object embedded within it, such as when the presentation of a kitchen scene facilitates a participant's ability to identify a loaf of bread, say (e.g., Biederman, 1972; Palmer, 1975; Davenport and Potter, 2004; though see Hollingworth and Henderson, 1998). Importantly, however, our environments typically convey contextual information via several different sensory modalities rather than just one. So, for example, when we are at the seaside, we perceive not only the blue sea and sky (hopefully), but also the sound of the waves crashing onto the beach, not to mention the smell of the salty sea air. Do such non-visual contextual cues also influence the visual perception of semantically related objects? In the present study, we investigated whether the semantic context provided by stimuli presented in another sensory modality (in this case, audition) modulate the perceptual outcome in vision; namely, visual awareness.

The phenomenon of binocular rivalry provides a fascinating window into human visual awareness (e.g., Crick, 1996). Binocular rivalry occurs when two dissimilar figures are presented to corresponding regions of the two eyes. Observers typically perceive one of the figures as dominant (while often being unaware of the presence of the other figure); after a while, the dominance of the figures may reverse and then keep alternating over time. This perceptual alternation has been attributed to the fact that the visual system receives ambiguous information from the two eyes and tries to find a unique perceptual solution, and therefore the information presented to each eye competes for control of the current conscious percept (see Alais and Blake, 2005, for a review). The fact that a constantly presented dichoptic figure induces alternating perceptual experiences in the binocular rivalry situation demonstrates the dynamic way in which the brain computes sensory information, a process that gives rise to a specific percept (e.g., Leopold and Logothetis, 1996).

Several researchers have tried to understand how visual awareness emerges in the binocular rivalry situation. According to an early view put forward by Helmholtz (1962), the alternation of perceptual dominance is under voluntary attentional control. Subsequently, researchers suggested that the phenomenon occurs as a result of competition between either two monocular channels (Levelt, 1965; Tong and Engel, 2001) or else between two pattern representations, one presented to each eye (Leopold and Logothetis, 1996; Logothetis et al., 1996; Tong et al., 1998). More recent models (e.g., Tong et al., 2006) have suggested that the mechanisms underlying binocular rivalry include not only competition at multiple levels of information processing (for reviews, see Tong, 2001; Blake and Logothetis, 2002), but also some form of excitatory connections that facilitate the perceptual grouping of visual stimuli (Kovacs et al., 1996; Alais and Blake, 1999), as well as top-down feedback, including attentional control and mental imagery (Meng and Tong, 2004; Mitchell et al., 2004; Chong et al., 2005;

van Ee et al., 2005; Pearson et al., 2008). That said, the underlying mechanisms giving rise to conscious perception in the binocular rivalry situation, while starting from interocular suppression, extend to a variety of different neural structures throughout the visual processing hierarchy.

Given that the phenomenon of binocular rivalry is, by definition, visual, one might have expected that the perceptual outcome for ambiguous visual inputs should thus be generated entirely within the visual system (cf. Hupé et al., 2008). On the other hand, however, some researchers have started to investigate whether visual awareness can be modulated by the information presented in another sensory modality. So, for example, it has recently been demonstrated that concurrently presented auditory cues can help to maintain the awareness of visual stimuli (Sheth and Shimojo, 2004; Chen and Yeh, 2008). Similar evidence has emerged from a binocular rivalry study demonstrating that the dominance duration of a looming (or rotating) visual pattern can be extended temporally when the rate of change of the visual stimulus happens to be synchronous with a series of pure tones or vibrotactile stimuli (or their combination, see van Ee et al., 2009). In addition, the directional information provided by the auditory modality can enhance the dominance duration of the moving random-dot kinematogram which happens to be moving in the same direction (Conrad et al., 2010).

Considering the seaside example outlined earlier, the meaning of a background sound (or soundtrack) plausibly provides a contextual effect on human information processing, which may, as a result, modulate the perceptual outcome that a person is aware of visually. Semantic congruency, which relies on the associations picked-up in daily life, provides an abstract constraint other than physical consistency between visual and auditory stimuli (such as coincidence in time or direction of motion mentioned earlier). This high-level factor has started to capture the attention of researchers interested in multisensory information processing (e.g., Greene et al., 2001; Molholm et al., 2004; van Atteveldt et al., 2004; Taylor et al., 2006; Iordanescu et al., 2008; Noppeney et al., 2008; Schneider et al., 2008; Chen and Spence, 2010; for a recent review, see Spence, 2011). On the other hand, modulations resulting from the presentation of semantically meaningful information have recently been documented by researchers studying unimodal binocular vision (Jiang et al., 2007; Costello et al., 2009; Ozkan and Braunstein, 2009). In the present study, we therefore investigated whether the semantic context provided by an auditory soundtrack would modulate human visual perception in the binocular rivalry situation.

Our first experiment was designed to test the crossmodal semantic modulatory effect on the dominant percept under conditions of binocular rivalry, while attempting to minimize or control any possible response biases elicited by the meaning of the sound. After first establishing this crossmodal effect, we then go on to explore the ways in which auditory semantic context modulates visual awareness in the binocular rivalry situation. Two visual factors, one high-level (selective attention) and one low-level (stimulus contrast) which have been shown to modulate visual perception in the binocular rivalry situation (Meng and Tong, 2004), are used to probe behaviorally the underlying mechanisms by which the auditory semantic context modulating visual awareness occurred in terms of current models of binocular rivalry (Tong et al., 2006).

## EXPERIMENT 1

In our first experiment, we investigated whether the semantic context of a background soundtrack would modulate the dominance of two competing percepts under the condition of binocular rivalry. The participants viewed a dichoptic figure consisting of a bird and a car (see **Figure 1**) while listening to a soundtrack. When studying audiovisual semantic congruency effects, the possibility that participants' responses are based on their utilizing a strategy designed to satisfy a particular laboratory task has to be avoided (see de Gelder and Bertelson, 2003). That is, there is a danger that the participants might merely report the stimulus that happened to be semantically congruent with the soundtrack rather than the percept that happened to be more salient (or dominant). In order to reduce the likelihood that the above-mentioned response bias would affect participants' performance, a novel experimental design was used in Experiment 1A: the participants only had to press keys to indicate the start and the end time of the perceptual dominance of the pre-designated figure (e.g., "bird") during the test period, while they listened to either the soundtrack that was *incongruent* with the visual target (i.e., a car soundtrack, in this case) or the sound that was *irrelevant* to both figures (i.e., a soundtrack recorded in a restaurant). A parallel task in which the pre-designated target figure was the car was also conducted. The participants listened to either the bird soundtrack in the incongruent condition or to the restaurant soundtrack in the irrelevant condition. Thus, the soundtrack was never congruent with the visual target that participants had to report. This aspect of the experimental design was introduced in order to reduce the likelihood that participants would simply report their perceptual dominance in accordance with whichever soundtrack they happened to hear. On the other hand, if the auditory semantic context can either prolong the dominance of the visual percept that happens to be semantically congruent with the soundtrack, or else shorten the dominance duration of the percept that happens to be semantically incongruent with the soundtrack, in the binocular rivalry situation, the dominance duration of the visual

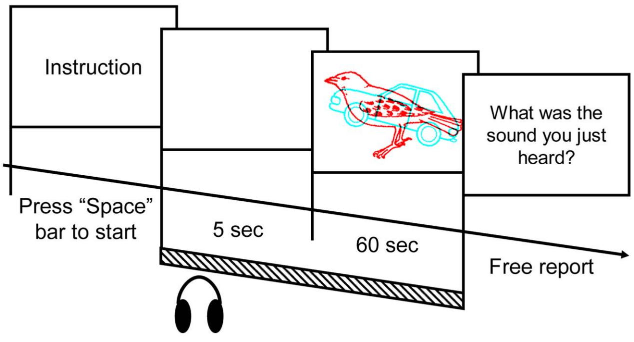

**FIGURE 1 | The trial sequence in Experiment 1A.** An example of the dichoptic stimulus pairs used in the present study is demonstrated in the third frame. The soundtrack was presented from the start of the blank fame until the end of the visual stimuli (i.e., for a total of 65 s).

target should be shorter in the *incongruent* than in the *irrelevant* condition.

## EXPERIMENT 1A

### *Participants*

Twelve volunteers (including the first author, three males, with a mean age of 26 years old) took part in this experiment in exchange for a £10 (UK Sterling) gift voucher or course credit. The other 11 participants were naïve as to the specific purpose of the study. They all had normal or corrected-to-normal vision and normal hearing by self report. The participants were tested using depth-defined figures embedded in red–green random-dot stereograms to ensure that they had normal binocular vision. The study has been approved by the ethic committee and human participant recruit system in Department of Experimental Psychology, University of Oxford. All of the participants were informed of their rights in accordance with the ethical standards laid down in the 1990 Declaration of Helsinki and signed a consent form.

### *Apparatus and stimuli*

The visual stimuli were presented on a 15′ color CRT monitor (75 Hz refresh rate). The participants sat at a viewing distance of 58 cm from the monitor in a dimly lit experimental chamber. The visual test stimuli consisted of the outline-drawings of a bird (4.44° × 2.76°) and car (4.41° × 2.27°) taken from Bates et al. (2003). The two figures were spatially superimposed, with the bird presented in red [CIE (0.621, 0.341)] and the car in cyan [CIE (0.220, 0.347)], or vice versa, against a white background [CIE (0.293, 0.332)]. These two color versions of the visual pictures (bird in red and car in cyan, or bird in cyan and car in red) were used to balance the influence of participants' dominant eye when viewing dichoptic figures. The participants wore glasses with a red filter on the left eye and a cyan filter on the right eye during the course of the experiment.

Three sound files, *bird* (consisting of birds singing in a forest), *car* (consisting of car horn and engine-revving sounds in a busy street), and *restaurant* (consisting of the sound of tableware clattering together in a restaurant), which had been recorded in realistic environments (downloaded from www.soundsnap.com on 06/11/2008) were used as the auditory soundtracks. The sound files were edited so that the auditory stimulus started from the beginning of the sound file and lasted for 65 s. The sounds were presented over closed-ear headphones and ranged in loudness from 55 to 68 dB SPL.

### *Design and procedure*

Two factors, semantic congruency (incongruent or irrelevant) and visual target (bird or car), were manipulated. Each participant reported the dominance of either the bird or the car percept in separate sessions in a counterbalanced order. Under those conditions in which the visual target was the bird, the participants were instructed to press the "1" key as soon as the image of the bird became dominant. The participants were informed that the criterion for responding that the bird was dominant was that they were able to see every detail, such as the texture of the wings, of the figure of the bird. As soon as any part of the bird figure became vague or else started to be occupied by the features of the car figure, they had to press the "0" key as soon as possible, to indicate that the image of the bird was no longer completely dominant. This criterion enabled us to estimate the dominance duration of the bird percept more conservatively, since it excluded those periods of time when the car percept being dominant as well as when participants experienced a mixed percept. Similarly, under those conditions in which the visual target was the car, the participants had to press "1" and "0" to indicate when they started and stopped perceiving the car percept as being dominant.

The participants initiated each trial by pressing the "SPACE" bar. A blank screen was presented for 5 s, followed by the presentation of the dichoptic figures for a further 60 s. The participants were instructed to fixate the area of the bird's wing and car door and to start reporting the dominance of the target figure as soon as the dichoptic figures were presented. They had to monitor the dominance of the target picture continuously during the test period. The participants were also instructed to pay attention to the context of the sound as well (in order to ensure that the soundtracks were processed; see van Ee et al., 2009, Experiment 4). At the end of the trial, the question "What sound did you just hear?" was presented on the monitor, and the participants had to enter their answer (free report) using the keyboard. The sound was presented from the onset of the blank frame until the offset of the visual stimuli, in order to allow participants sufficient time to realize what the semantic context conveyed by the soundtrack was.

In both visual tasks (i.e., when the visual target was a bird and when it was a car), a block of 12 trials was presented (consisting of two sound conditions × two color versions of visual pictures, each conditions were repeatedly tested three times). The order of presentation of these 12 trials was randomized. Prior to the completion of the experimental block of trials, a practice block containing six no-sound trials was presented in order to familiarize the participants with the task. The participants were instructed to establish their criterion for reporting the exclusive dominance of the target picture, and to try and hold this criterion constant throughout the experiment. The experiment lasted for approximately 1 h.

### *Results*

The *proportion of time for which the target percept was dominant* was calculated by dividing the sum of each dominance duration of the target percept by 60 s. Note that the participants may have occasionally pressed the "1" or "0" key twice. In such cases, the shorter duration (i.e., the duration from the second "1" keypress to the first "0" keypress) was used.

A two-way analysis of variance (ANOVA) was conducted with the factors of semantic congruency (incongruent or irrelevant) and visual target (bird or car; see **Figure 2A**)[1]. The results revealed significant main effects of both semantic congruency

---

[1] For reasons that are unknown, one participant stopped reporting the dominance of the target figure during the first 25 s in a trial in the incongruent condition, while he/she kept reporting the dominance percept until the end in all of the other trials. This trial, as well as a matched color version trial in the irrelevant condition, was excluded from the data analysis. By doing this, we were able to ensure that the data in the incongruent and irrelevant conditions came from equal number of trials from the two color versions, so that the factor of participants' eye dominance could be matched.

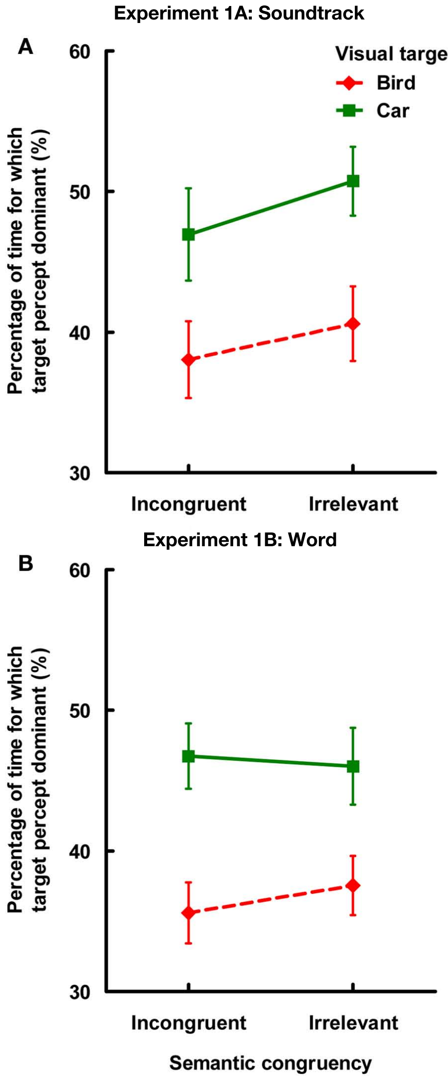

**FIGURE 2 | Results of the proportion of time for which the target percept was dominant (either bird or car, proportion of the 60-s viewing period) in Experiments 1A and 1B [(A,B), respectively].** Error bars represent ±1 SE of the mean.

[$F(1,11) = 25.68$, MSE = 0.0005, $p < 0.0005$, $\eta_p^2 = 0.71$] and visual target [$F(1,11) = 11.50$, MSE = 0.01, $p < 0.01$, $\eta_p^2 = 0.51$]. There was, however, no interaction between these two factors [$F(1,11) = 0.60$, MSE = 0.001, $p = 0.46$, $\eta_p^2 = 0.07$]. The planned simple main effect of the semantic congruency factor revealed that the proportion of dominance of the target picture was lower when listening to the *incongruent* soundtrack than when listening to the *irrelevant* soundtrack both when the visual target was the bird [$F(1,22) = 6.34$, MSE = 0.001, $p < 0.05$, $\eta_p^2 = 0.15$], as well as when it was the car [$F(1,22) = 13.96$, $p < 0.005$, $\eta_p^2 = 0.33$]. In addition, the magnitude of the auditory modulatory effect (incongruent vs. irrelevant) was not significantly different in the bird and car target conditions [$F(1,11) = 0.60$, MSE = 0.002, $p = 0.45$, $\eta_p^2 = 0.06$].

## EXPERIMENT 1B

Two further possibilities regarding the crossmodal semantic modulation reported in Experiment 1A need to be considered. First, the presented soundtrack may have accessed its associated abstract semantic representation and then modulated the dominant percept in the binocular rivalry situation. In this case, the semantic modulation constitutes a form of top-down semantic modulation rather than a form of audiovisual interaction. Second, even though the design of Experiment 1A effectively avoids the bias that the participants strategically reported the percept that is congruent with the meaning of the soundtrack as being dominant, it is important to note that a second type of bias should also be considered. That is, it could be argued that the presentation of the incongruent soundtrack may have provided a cue that discouraged the participants from reporting the target percept as being dominant, as compared to the presentation of the irrelevant soundtrack.

Experiment 1B was designed to control for the possibility that the crossmodal semantic modulation effects observed thus far might simply have resulted from the participants holding an abstract concept in mind, as well as the response bias elicited by the presentation of a cue that was incongruent with the identity of the visual target. Rather than presenting a soundtrack, the name of one of the soundtracks was presented on the monitor for 5 s prior to the presentation of the dichoptic figures (during this period, a blank frame had been presented in Experiment 1A). That is, the participants were provided with a word (the associated name of the soundtracks used in Experiment 1A) that was either incongruent with or irrelevant to the visual target, while they were tested in silence during the subsequent test period. The participants were instructed to retain the word in memory and to report it at the end of each trial, in order to ensure that they had maintained this semantic cue during the course of the test period. The word therefore provided an abstract semantic cue to the participants. In addition, the presentation and retention of this semantic cue in memory by participants would be expected to elicit a similar response bias in the incongruent (as compared to the irrelevant) condition. Our prediction was that if an abstract semantic cue or the response bias elicited by the incongruent cue (rather than the audiovisual semantic interaction) was sufficient to induce the semantic effect in the binocular rivalry situation, the significant difference between incongruent and irrelevant conditions should still be observed.

Two factors, semantic congruency (incongruent or irrelevant) and visual target (bird or car), were manipulated in this experiment. When the visual target was the bird, the words "car" and "restaurant" were presented in the incongruent and irrelevant conditions, respectively. Similarly, when the visual target was the car, the words "bird" and "restaurant" were presented in the incongruent and irrelevant conditions, respectively. The other experimental details were exactly the same as in Experiment 1A.

### *Participants*

Twelve volunteers (including the first author, six males, with a mean age of 24 years old) took part in this experiment. The other 11 participants did not attend Experiment 1A and they were naïve as to the goal of the study. The other details are the same as in Experiment 1A.

#### *Results*

The data were analyzed in the same manner as in Experiment 1A. The participants misreported the word in six trials (out of total 288 trials). These trials, as well as matched color version trials in the other word condition, were excluded from the analysis (4.2% of total trials). A two-way ANOVA was conducted with the factors of semantic congruency (incongruent or irrelevant) and visual target (bird or car; see **Figure 2B**). Once again, the results revealed a significant main effect of visual target [$F(1,11) = 7.99$, $\mathrm{MSE} = 0.01$, $p < 0.05$, $\eta_p^2 = 0.42$]. Critically, however, neither the main effect of semantic congruency [$F(1,11) = 0.96$, $\mathrm{MSE} = 0.0005$, $p = 0.35$, $\eta_p^2 = 0.08$], nor the interaction term [$F(1,11) = 1.49$, $\mathrm{MSE} = 0.001$, $p = 0.25$, $\eta_p^2 = 0.11$] was significant. The planned simple main effect of the semantic congruency factor revealed that the proportion of dominance of the target picture was *not* significant as a function of whether the condition was incongruent or irrelevant when the visual target was either the bird [$F(1,22) = 2.39$, $\mathrm{MSE} = 0.001$, $p = 0.15$, $\eta_p^2 = 0.09$], or the car [$F(1,22) = 0.33$, $p = 0.57$, $\eta_p^2 = 0.01$].

#### *Comparison of Experiments 1A and 1B*

In order to verify that the semantic modulation on the dominant percept in the binocular rivalry situation was significant in Experiment 1A but not in Experiment 1B, a three-way ANOVA on the factor of cue type (soundtrack or word), semantic congruency (incongruent or irrelevant), and visual target (bird or car) was conducted. The between-participants factor was cue type while the latter two factors were varied on a within-participants basis. The results revealed significant main effects of semantic congruency [$F(1,22) = 18.47$, $\mathrm{MSE} = 0.0005$, $p < 0.0005$, $\eta_p^2 = 0.47$] and visual target [$F(1,22) = 18.75$, $\mathrm{MSE} = 0.01$, $p < 0.0005$, $\eta_p^2 = 0.46$]. Critically, the interaction between cue type and semantic congruency was significant [$F(1,22) = 8.43$, $\mathrm{MSE} = 0.0005$, $p < 0.01$, $\eta_p^2 = 0.29$]. The simple main effect of the semantic congruency factor was significant when the cue was a soundtrack [$F(1,22) = 25.93$, $\mathrm{MSE} = 0.0005$, $p < 0.0001$, $\eta_p^2 = 0.55$], but not when the cue was a word [$F(1,22) = 0.97$, $p = 0.34$, $\eta_p^2 = 0.04$]. The magnitude of the auditory modulatory effect (incongruent vs. irrelevant) was submitted to a two-way ANOVA on the factor of cue type and visual target. Only the main effect of cue type reached significance [3.2 vs. 0.6% for soundtrack and word condition, respectively, $F(1,22) = 8.39$, $\mathrm{MSE} = 0.001$, $p < 0.01$, $\eta_p^2 = 0.28$]. The other main effect of visual target [$F(1,22) = 0.26$, $\mathrm{MSE} = 0.002$, $p = 0.61$, $\eta_p^2 = 0.02$] and the interaction term [$F(1,22) = 2.06$, $\mathrm{MSE} = 0.002$, $p = 0.17$, $\eta_p^2 = 0.09$] were not significant.

#### *Discussion*

The results of Experiment 1A therefore demonstrate a crossmodal modulation on the proportion of dominance measure resulting from the auditory semantic context that was present in the binocular rivalry situation. Note that the soundtrack to which the participants listened during the test period was *never* congruent with the visual target. That is, the crossmodal semantic modulation of binocular rivalry was *indirect* in terms of the meaning of the sound either increasing the amount of time for which the participant perceived the non-target visual picture and/or decreasing the amount of time for which they perceived the target picture. Both possibilities would have led to a reduction in the proportion of dominance of the target picture. This feature of the design means that it was not the case that the meaning of the sound directly biased the participants' response to report the semantically congruent visual stimulus as being dominant.

On the other hand, the results in Experiment 1B, demonstrated that simply maintaining a word in memory during the test period did not bias the participants' visual perception or responses. Note that the comparison of the results of Experiments 1A and 1B is meaningful based on the *a priori* assumption that the presentation of a semantically congruent (though task-irrelevant) word can prime the participants' performance regarding the picture (e.g., Glaser and Glaser, 1989). Hence, the modulatory effect of auditory semantic content reported in Experiment 1A cannot simply be attributed to a semantic priming effect elicited by activating an abstract concept regarding one of the pictures (cf. Balcetis and Dale, 2007), nor to any response bias that was potentially elicited by the presentation of a soundtrack that was incongruent with the visual target (i.e., congruent with the competing percept). In addition, due to the fact that the participants were continuously receiving the auditory information during the test period in Experiment 1A while simply provided a semantic cue before the test period in Experiment 1B, we suggest that the crossmodal semantic congruency effect should be perceptual in nature (i.e., depending on the input of sensory information) rather than simply a conceptual effect (depending on the prior acquired knowledge). These results therefore highlight a significant crossmodal modulation of perceptual dominance in the binocular rivalry situation. This result can be attributed to the semantic context embedded in the auditory soundtrack that the observers were listening to.

In both experiments, the results revealed that the proportion of dominance was larger when the car was the target than when the bird was the target. Note that the wing of the bird contained small individual elements constituting the texture of a feather, so the individual elements may have disappeared occasionally (Kovacs et al., 1996). On the other hand, most of the lines making up the figure of the car were connected, and so they should group into a unitary element (such as a car door). Since the participants were instructed to report the target figure as being dominant *only* when they could see all of its features, the well-grouped figure (i.e., the car) should have reached this criterion more easily than the less-well-grouped figure (i.e., the bird).

In van Ee et al.'s (2009) study, it was reported that an auditory stimulus enhanced the dominance duration of a synchronous visual pattern only when participants happened to attend to that visual pattern in the binocular rivalry situation (see their Experiment 1). In Experiment 1A of the present study, given that the participants had to simply monitor the dominance of one of the two figures, their goal-directed attention should presumably have been focused voluntarily (i.e., endogenously) on the target stimulus. Nevertheless, in order to further investigate the interplay between crossmodal semantic congruency and the participants' selective attention on the perception of binocular rivalry, these two factors were manipulated independently in Experiment 2A. In addition, given that the visual competition mechanism in binocular rivalry started from a low-level interocular suppression, binocular rivalry is presumably susceptible to stimulus saliency (Mueller and Blake, 1989). We therefore decided to test the interplay between crossmodal semantic congruency and the manipulation of the stimulus contrast in Experiment 2B.

## EXPERIMENT 2

We designed two further experiments in order to measure whether the crossmodal semantic congruency effect is robust when simultaneously manipulating visual factors that have previously been shown to modulate participants' perception in the binocular rivalry situation, such as participants directing their selective attention to a specific percept or increasing the stimulus contrast of one of the dichoptic images (Meng and Tong, 2004; van Ee et al., 2005). Besides, knowing whether the modulations of crossmodal semantic congruency and either visual factor (i.e., selective attention or stimulus contrast) work additively or interactively to influence human visual perception would help us understand the possible mechanism underlying the crossmodal effect of auditory semantic context. We therefore manipulated auditory semantic congruency and visual selective attention in Experiment 2A, and auditory semantic congruency and visual stimulus contrast in Experiment 2B, respectively.

In Experiments 2A and 2B, the participants performed a typical binocular rivalry experiment reporting the percept that was subjectively dominant. That is, the participants had to press the "1" key whenever the image of the bird was dominant, and the "0" key whenever the image of the car was dominant. Three performance indices were used: the *proportion of time for which the bird percept was dominant* was calculated by dividing the sum duration of each bird percept by the sum duration of both the bird and car percept within the test period (thus, the proportion of time for the dominance of the bird and car views were reciprocally related). Accordingly, this measure would be expected to increase following any experimental manipulation that favored the bird percept (i.e., maintaining the bird percept or the bird figure by means of its higher contrast), whereas it should be decreased by the manipulation favoring the car percept (i.e., maintaining the car percept or the car figure by means of its higher contrast). The second index consisted of the average *number of switches* between the bird and car percept that took place during the test period in each sound condition. This index links closely to the idea of voluntary control in the binocular rivalry situation: in particular, when participants try to maintain a particular percept, they are able to delay the switch to the other percept (van Ee et al., 2005). It has been suggested that the combination of an increase in the proportion of dominance duration as well as a reduction in the number of switches can be considered as the signature of selective attention in the binocular rivalry situation (see van Ee et al., 2005). The third index was the number of times that the *first percept* was the bird out of six trials in each sound condition. This index can be considered as the result of initial competition between the images presented to each eye.

## EXPERIMENT 2A

In Experiment 2A, the target of participants' selective attention over the dichoptic figure was manipulated independently of the meaning of the sound. That is, the participants were instructed to maintain the bird percept, to maintain the car percept, or to view the figures passively in the control condition (see Meng and Tong, 2004; van Ee et al., 2005). Meanwhile, the participants either heard the birds singing or else the revving car engine soundtracks.

### *Participants*

Seven participants (including the first author, three males, with a mean age of 24 years old) took part in Experiment 2A. All of them had prior experience of binocular rivalry experiments, and three of them had taken part in Experiment 1A. However, the other six participants (except the author) were naïve regarding the goal of the present experiment.

### *Design and procedure*

Two factors, sound (bird and car) and selective attention (passive, maintain bird, or maintain car) were manipulated. The 5-s blank frame presented prior to the visual stimulus display now contained the instruction to "just look at the figures PASSIVELY," "try to maintain the percept of the BIRD as long as possible" or "try to maintain the percept of the CAR as long as possible." Note that in all three conditions, the participants had to report their current dominant percept (either bird or car). The bird or car soundtrack started at the onset of the attention instruction frame. The figures were larger than those used in Experiment 1 (bird: 7.85° × 6.39°; car: 8.34° × 3.95°).

Three blocks of experimental trials were presented. There were 12 trials (two sound conditions × three selective attention conditions × two color versions of the pictures) presented in a randomized order in each block. A practice block containing six no-sound trials, two for each selective attention condition, was conducted prior to the main experiment.

### *Results*

A two-way ANOVA was conducted on the factors of sound and selective attention for each index separately. In the analysis of the data concerning the proportion of dominance data (see **Figure 3A**), there were significant main effects of sound [$F(1,6) = 17.25$, MSE = 0.004, $p < 0.01$, $\eta_p^2 = 0.75$], and selective attention [$F(2,12) = 12.63$, MSE = 0.04, $p < 0.005$, $\eta_p^2 = 0.68$]. A *post hoc* Tukey's test revealed that there was a significant difference between the *maintain bird* and *maintain car* conditions ($p < 0.01$). The interaction between sound and selective attention was, however, not significant [$F(2,12) = 1.39$, MSE = 0.002, $p = 0.29$, $\eta_p^2 = 0.18$]. In order to examine whether the crossmodal semantic modulation of binocular rivalry was significant in all three selective attention conditions, planned simple main effects on the sound factor were conducted. The results revealed that the bird percept was dominant for a larger proportion of the time when presenting the bird soundtrack than when presenting the car soundtrack in the maintain bird condition [$F(1,18) = 5.57$, MSE = 0.002, $p < 0.05$, $\eta_p^2 = 0.13$], in the passive condition [$F(1,18) = 16.65$, $p < 0.001$, $\eta_p^2 = 0.37$], and in the maintain car condition [$F(1,18) = 4.91$, $p < 0.05$, $\eta_p^2 = 0.11$]. The magnitude of the auditory semantic modulation effect (bird- vs. car-sound) was not significantly different in the three selective attention conditions [$F(2,12) = 1.40$, MSE = 0.004, $p = 0.29$, $\eta_p^2 = 0.19$].

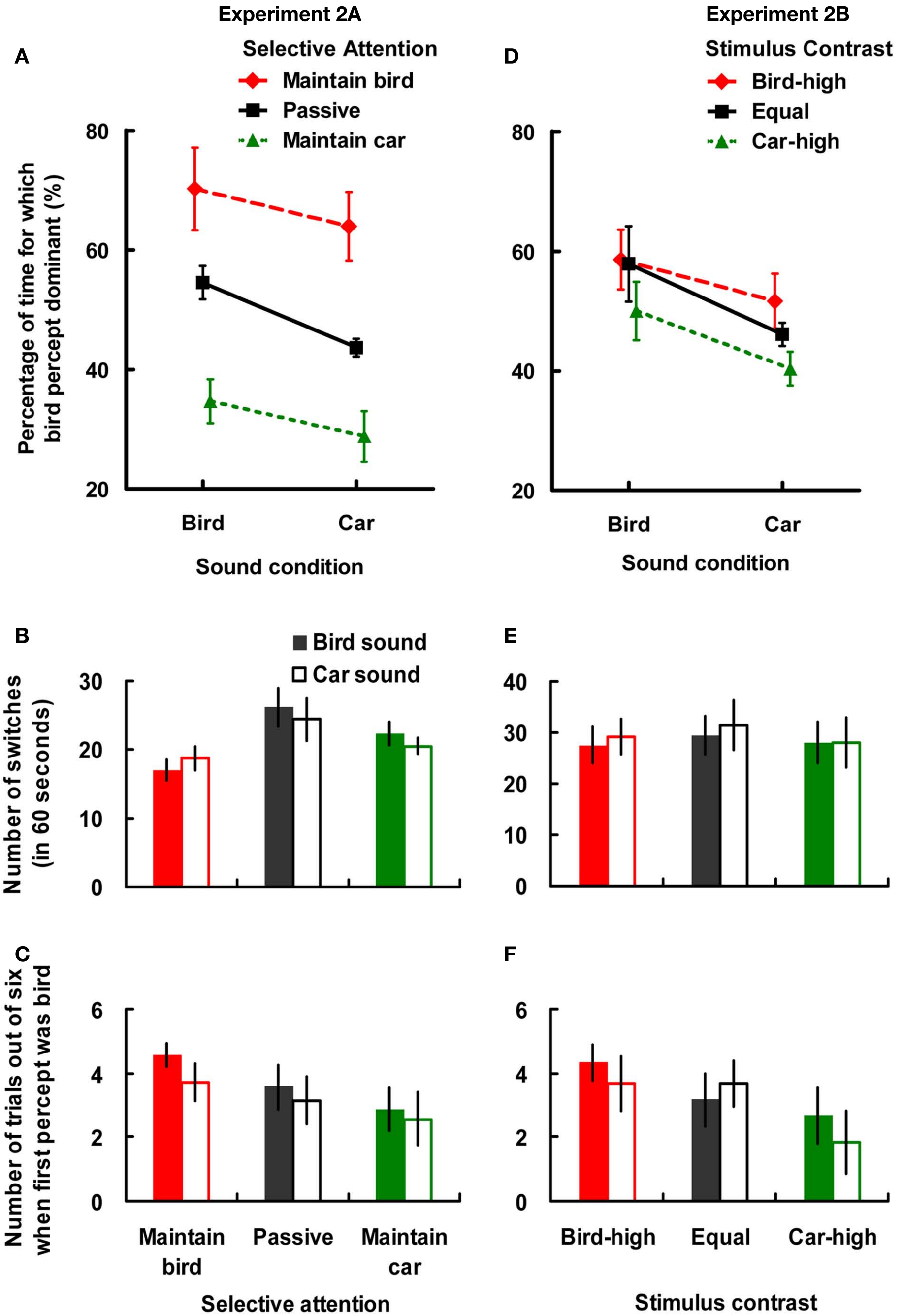

**FIGURE 3 | Results of the bird- and car-sound conditions under the manipulation of either selective attention (A–C) in Experiment 2A or stimulus contrast (D–F) in Experiment 2B in the binocular rivalry situation.** The upper panels **(A,D)** depict the proportion of time that the bird percept was dominant (proportion to total dominance duration of both bird and car percepts); the middle panels **(B,E)** show the average number of switches that took place in each trial (during the course of 60 s); the lower panels **(C,F)** show the number of times that participants reported the bird as the first percept (out of six trials). **(A)** The proportion of time that the bird percept was dominant in the maintain bird (dashed line), passive viewing (solid line), and maintain car (dotted line) conditions of Experiment 2A. **(D)** The proportion of time that the bird percept was dominant in the bird-high (dashed line), equal (solid line), and car-high (dotted line) conditions in Experiment 2B. Error bars represent ±1 SE of the mean.

Analysis of the number of switches in each condition (see **Figure 3B**) revealed significant main effects of both sound [$F(1,6) = 8.19$, MSE $= 0.43$, $p < 0.05$, $\eta_p^2 = 0.58$] and selective attention [$F(2,12) = 3.92$, MSE $= 48.05$, $p < 0.05$, $\eta_p^2 = 0.40$]. A *post hoc* Tukey's test revealed that the number of switches was higher in the passive condition than in the maintain bird condition ($p < 0.05$). The interaction between sound condition and selective attention was also significant [$F(2,12) = 5.27$, MSE $= 2.57$,

$p < 0.05$, $\eta_p^2 = 0.47$]. The simple main effect revealed that the number of switches was *smaller* in the *bird-sound* condition than in the *car-sound* condition when the participants had to try and maintain their view of the bird [$F(1,18) = 5.39$, MSE = 1.85, $p < 0.05$, $\eta_p^2 = 0.16$]. By contrast, the number of switches was *smaller* in the *car-sound* condition than in the *bird-sound* condition when the participants passively viewed the figures [$F(1,18) = 5.09$, $p < 0.05$, $\eta_p^2 = 0.15$], or when they had to try and maintain the view of the car [$F(1,18) = 6.02$, $p < 0.05$, $\eta_p^2 = 0.17$].

The results of the first percept in each condition (see **Figure 3C**) revealed that the participants reported the bird as the first percept somewhat more frequently when they were either listening to the bird sound or when trying to maintain the bird view, as compared to when they were either listening to the car sound or else trying to maintain their view of the car. However, neither of the main effects, nor their interaction, reached statistical significance (all $F\text{s} < 3.27$, $p\text{s} > 0.12$, $\eta_p^2 = 0.35$).

### *Discussion*

These results demonstrate that the participants reported the bird percept as being dominant for more of the time (i.e., the proportion of dominance was larger) when they heard the sound of birds singing than when they heard the sound of cars revving their engines. In addition, selective attention also modulated the proportion of dominance of the bird percept when the participants tried to maintain their view of the bird or car, consistent with the results reported by both Meng and Tong (2004) and van Ee et al. (2005). Nevertheless, the interaction between these two factors was not significant. Critically, the modulation of auditory semantic context was robustly observed in the maintain bird, passive, and maintain car conditions, and what is more, the magnitude of the crossmodal modulatory effect was similar in these three conditions. In other words, the crossmodal semantic modulation observed in Experiment 2A cannot be attributed solely to the meaning of the sound guiding participants' attentional selection, either voluntarily or involuntarily. If it had been the case, similar results in terms of the proportion of dominance measure should have been observed in the bird-sound/maintain bird and bird-sound/passive conditions because they both depend on the participants' devoting attention to the bird percept. In the same vein, similar results should have been observed in the car-sound/maintain car and car-sound/passive conditions because they both depend on the participants' devoting attention to the car percept. As a result, we should have observed that the effect of auditory semantic context was reduced or eliminated when the participants had to attend to a specific percept during the test period (see Hsiao et al., 2010).

On the other hand, the crossmodal semantic modulation of visual awareness in the binocular rivalry situation may be mediated (or enhanced) by selective attention as demonstrated by van Ee et al. (2009). These researchers reported a crossmodal modulation of binocular rivalry perception by following the presentation of series of beeps. However, this effect was only observed when the participants happened to attend to the temporally-coincident percept rather than when they passively viewed the dichoptic figure (Experiment 1 in their study). By contrast, we observed the crossmodal semantic congruency effect on the participants' proportion of dominance measure in the passive condition, while the magnitude was *no* larger in the maintain bird and maintain car conditions. Nevertheless, a possible explanation for this result is that this measure had almost reached ceiling (or floor) in the bird-sound/maintain bird and the car-sound/maintain car conditions.

In terms of the number of switches, the results reveal that selective attention effectively reduced the frequency of perceptual switches during the test period (see also van Ee et al., 2005). This result is in line with the fact that the occurrence of switches from one percept to the other under conditions of binocular rivalry can be modulated by attention (see Lumer et al., 1998; Leopold and Logothetis, 1999). We therefore observed that selective attention modulated both the proportion of dominance and number of switches measures (van Ee et al., 2005). Note that the attentional effect was more obvious in the maintain bird condition than in the maintain car condition. This result may have been due to the fact that it is harder for the bird percept to dominate (see Experiment 1). Consequently, more attentional effort should be devoted in the maintain bird condition to maintain it. It should, however, also be noted that the auditory semantic context reported in Experiment 2A somehow assisted visual attentional control over perceptual switching. This is evidenced by the fact that when the participants were listening to the bird soundtrack and were instructed to try and maintain the bird percept in awareness, the number of switches was smaller than when they heard the car soundtrack (see the opposite patterns modulated by the sound in the maintain bird and the other two conditions). It is possible that since attentional control over a given object representation relies on holding that target in working memory (see Desimone and Duncan, 1995), the presentation of a semantically congruent auditory soundtrack may have helped the participants to hold the target in mind during the test period.

In summary, the crossmodal semantic modulation was robust in the binocular rivalry situation irrespective of the participants' state of selective attention in terms of determining the proportion of dominance measure. We therefore suggest that attentional control over a specific percept is not a necessary condition for the crossmodal modulation by auditory semantic context in the binocular rivalry situation (cf. van Ee et al., 2009). In addition, we also observed that crossmodal semantic modulation and visual attentional control interacted in terms of the switch times measure.

## EXPERIMENT 2B

The final experiment in the present study addressed the question of whether the modulatory effect of auditory semantic context would interact with low-level visual factors in determining the consequences of binocular rivalry. The luminance contrast of a figure provides a bottom-up (i.e., stimulus-driven) factor. That is, a higher luminance contrast figure will likely win the initial competition and be perceived for more of the time than a figure with a lower luminance contrast (Mueller and Blake, 1989; Meng and Tong, 2004). The participants in this experiment heard either the bird or car soundtrack while presented with one of three levels of luminance contrast (see below).

### *Participants*

Six of the participants who took part in Experiment 2A (one dropped out) were tested.

### Design and procedure

Two factors, sound (bird and car) and stimulus contrast (equal, bird-high, car-high) were manipulated. In the equal condition, the dichoptic figures used in the previous experiments were presented. In the bird-high condition, the luminance contrast of the bird figure was constant (the Michelson Contrast value measured through the color filter was 85.4% for red and 70.7% for cyan), whereas the luminance contrast of the car figure was reduced (the Michelson Contrast value was 81.0% for red and 65.9% for cyan), and vice versa in the car-high condition. In each trial, the frame that normally provides the attentional instruction was now left blank (just as in Experiment 1A). The participants were instructed to view the figures passively.

Three blocks of experimental trials were presented. There were 12 trials (two sound conditions × three stimulus contrast conditions × two color versions of visual pictures) presented in a randomized order in each block. A practice block containing six no-sound trials, two for each stimulus contrast condition, was conducted before the main experiment. The other details were the same as Experiment 2A.

### Results

A two-way ANOVA was conducted with the factors of sound and stimulus contrast for each index separately. Analysis of the proportion of dominance data (see **Figure 3D**) revealed significant main effects of sound [$F(1,5) = 7.31$, $\mathrm{MSE} = 0.01$, $p < 0.05$, $\eta_p^2 = 0.59$] and stimulus contrast [$F(2,10) = 11.76$, $\mathrm{MSE} = 0.003$, $p < 0.005$, $\eta_p^2 = 0.71$]. A *post hoc* Tukey's test revealed that there was a significant difference between the bird-high and car-high conditions ($p < 0.01$), and between the equal and car-high conditions ($p < 0.05$). The interaction between sound and stimulus contrast was, however, not significant [$F(2,10) = 1.14$, $\mathrm{MSE} = 0.002$, $p = 0.36$, $\eta_p^2 = 0.20$]. The planned simple main effect of the sound factor revealed that the proportion of dominance of the bird percept was higher when participants heard the bird soundtrack than when they heard the car soundtrack in the equal luminance condition [$F(1,15) = 8.82$, $\mathrm{MSE} = 0.005$, $p < 0.01$, $\eta_p^2 = 0.27$], in the car-high condition [$F(1,15) = 5.97$, $p < 0.05$, $\eta_p^2 = 0.18$], while failing to reach significance in the bird-high condition [$F(1,15) = 3.03$, $p = 0.10$, $\eta_p^2 = 0.09$][2]. Note that the magnitude of the auditory modulation effect (bird- vs. car-sound) was not significantly different across the three levels of stimulus contrast [$F(2,10) = 1.12$, $\mathrm{MSE} = 0.003$, $p = 0.37$, $\eta_p^2 = 0.18$]. No significant differences were observed in the analysis of the number of switches (see **Figure 3E**); all $F$s $< 4.01$, $p$s $> 0.05$, $\eta_p^2 < 0.44$.

Analysis of the first percept data (see **Figure 3F**) revealed a significant main effect of stimulus contrast [$F(2,10) = 7.18$, $\mathrm{MSE} = 1.33$, $p < 0.05$, $\eta_p^2 = 0.59$]. A *post hoc* Tukey's test revealed a significant difference between the bird-high and car-high conditions ($p < 0.05$). The other main effect, that of sound [$F(1,5) = 0.83$, $\mathrm{MSE} = 1.20$, $p = 0.40$, $\eta_p^2 = 0.14$], and the interaction between these two factors [$F(2,10) = 2.02$, $\mathrm{MSE} = 0.78$, $p = 0.18$, $\eta_p^2 = 0.29$], were not significant.

### Discussion

The results of Experiment 2B once again demonstrate that auditory semantic context can increase the proportion of dominance of a semantically congruent percept under conditions of binocular rivalry. Besides, we also replicated the finding that the proportion of dominance is influenced by the relative luminance contrast of the two visual figures (Meng and Tong, 2004). Critically, there was no interaction between the modulation by sound and stimulus contrast on the proportion of dominance measure. This result also indicates that even when the participants were listening to the bird (or car) soundtrack, they were still sensitive to the low-level visual properties (i.e., stimulus contrast in this experiment) of the dichoptic figures during the test period.

On the other hand, the results of Experiment 2B reveal that only the stimulus contrast determined the first percept whereas the auditory semantic context did not (see also Experiment 2A; though see Rommetveit et al., 1968; Costello et al., 2009). That is, the figure that had the higher contrast was perceived first. It should be noted that the measurement of which picture reached awareness first (i.e., the first percept) may merely reflect the result of dichoptic masking rather than genuine binocular rivalry (as indexed by the proportion of dominance duration of a given percept during the test period, see Blake, 1988, p. 140; Noest et al., 2007). That said, the results of Experiment 2B revealed that both visual stimulus contrast and auditory semantic context can modulate the perceptual outcome of binocular rivalry (in an additive fashion), while the former was more dominant than the latter in terms of determining the perceptual outcome of dichoptic masking. We therefore suggest that these two factors, visual stimulus contrast and auditory semantic context, can be dissociated in terms of both the proportion of dominance and first percept measures. On the other hand, the results reported here also suggest that, even though dichoptic masking and binocular rivalry may involve a similar mechanism of interocular suppression (since they were both sensitive to visual stimulus contrast), binocular rivalry seems to involve the later stages of visual processing (perhaps including the semantic level) as well (see Noest et al., 2007; van Boxtel et al., 2007; Baker and Graf, 2009).

## GENERAL DISCUSSION

The results of the two experiments reported in the present study demonstrate that a participant's visual awareness in the binocular rivalry situation can be modulated by the semantic context provided by a concurrently presented auditory soundtrack. In Experiment 1A, the proportion of dominance measure of the target percept was smaller when the participant listened to a soundtrack that was incongruent (i.e., that was congruent with the competing percept) than to a soundtrack that was irrelevant to both percepts (i.e., the restaurant soundtrack in the present study). Besides, our results also highlighted the fact that the proportion of dominance measure was unaffected by the instruction to maintain a word in working memory (rather than continuously hearing a soundtrack) during the test period (Experiment 1B). In terms of the proportion of dominance results, we further demonstrated

---

[2] It should be noted that in the bird-high condition, all six participants consistently reported the bird percept for more of the time when listening to the bird sound than when listening to the car sound. A repeated measure $t$-test revealed this difference to be significant [$t(5) = 2.22$, $p < 0.05$, one-tailed].

that the modulation by auditory semantic context was additive with that resulting from visual selective attention (Experiment 2A) and additive with that resulting from visual luminance contrast as well (Experiment 2B). Each of the three factors, however, may influence other aspects of participants' performance in the binocular rivalry situation. So, for example, visual selective attention effectively reduced the switch times during the test period (see Experiment 2A; see also van Ee et al., 2005). On the other hand, visual stimulus contrast effectively modulated the participants' first percept (Experiment 2B). Note that the auditory semantic context only modulated switch times when it was simultaneously manipulated with visual selective attention (in Experiment 2A, but not in Experiment 2B), and never modulated the first percept. Considering all these three indices allows us to suggest that the crossmodal modulation of which by auditory semantic context can be dissociated from that by visual selective attention and visual stimulus contrast to a certain extent.

One of the more novel observations to emerge from the results reported here is that the semantic context of an auditory soundtrack can effectively modulate participants' visual awareness. As an extension to the previous studies reported by Sheth and Shimojo (2004), van Ee et al. (2009), and Conrad et al. (2010) in which the visibility of stimuli undergoing visual competition was maintained by the presentation of an auditory stimulus in terms of its physical properties (such as temporal synchrony or direction of motion), the results reported here demonstrate that the crossmodal modulation of binocular rivalry by sound can also extend to its semantic context. The results of the present study therefore provide important evidence that the factors modulating binocular rivalry can reach the semantic level (Engel, 1956; Yang and Yeh, 2011; though see Zimba and Blake, 1983; Blake, 1988), and critically, can occur crossmodally.

It should be noted that since our participants were instructed to attend to the soundtrack during the test period, the soundtrack may be not automatically processed (cf. Schneider et al., 1984; see also van Ee et al., 2009). Here we would rather suggest that the soundtrack should be "selected in," rather than simply "filtered out," in the early stages of auditory information processing (e.g., Treisman and Riley, 1969). Nevertheless, it is possible that once the sound had been processed, the auditory semantic context then unavoidably interacts with any relevant visual information (see Treisman and Davies, 1973; Brand-D'Abrescia and Lavie, 2008; Yuval-Greenberg and Deouell, 2009).

Recently, Pearson et al. (2008) demonstrated that generating the visual mental image of the percepts in a binocular rivalry situation can increase the possibility of that percept winning the competition to reach awareness. Can the crossmodal semantic modulation reported here have been the result of participants generating a visual mental image corresponding to the soundtrack that they happened to be listening to? Mental imagery can be considered as providing a top-down means of modulating a particular object representation, though the time required to generate a mental image is much longer than that required to execute a shift of selective attention (see Pearson et al., 2008, Experiment 3). Note, however, that Pearson et al. (2008) also reported that when the background was 100% illuminated (i.e., a white background, as in the present study), their participants performed similarly under the conditions of viewing passively and generating a mental image. Besides, Segal and Fusella's (1969, 1970) early studies demonstrate that a person's sensitivity to detect a visual (or auditory) target was lowered when he/she imaged that stimulus in the same sensory modality. Such modality-specific suppression during mental image generation has recently been observed in primary sensory areas in humans (i.e., visual and auditory cortices, see Daselaar et al., 2010). It is true that in most of the experiments reported by Pearson et al. (2008) there was no visual stimulus presented during the imagery period. Therefore, it seems unlikely that a person's ability to invoke mental imagery can be used to enhance a particular percept in the binocular rivalry situation where the visual background was white and visual and auditory stimuli were continuously presented, as in the present study.

## POSSIBLE MECHANISMS OF THE MODULATION BY AUDITORY SEMANTIC CONTEXT ON VISUAL AWARENESS

The results of Experiments 2A and 2B demonstrated that auditory semantic context, visual selective attention, and visual stimulus contrast, all modulated participants' visual perception under conditions of binocular rivalry. All three factors effectively modulated the typical index of proportion of dominance duration of a given percept. Considering the fact that three indices we used, the factors of auditory semantic context and visual stimulus contrast can be dissociated; however, while the modulation of auditory semantic context was not necessarily mediated by visual selective attention, these two factors may interact to some degree.

Let us then consider how the audiovisual semantic congruency effect reported here could be implemented in the model of binocular rivalry based on the three mechanisms (inhibitory, lateral excitatory, and feedback connections) proposed by Tong et al. (2006; see **Figure 4**). The modulation by visual stimulus contrast can be accounted for by interocular inhibition (Tong, 2001), whereas the modulation elicited by visual selective attention can be accounted for by feedback connections (Tong et al., 2006). Auditory semantic context likely enhanced the representation of semantically congruent visual object representation (Iordanescu et al., 2008; Chen and Spence, 2010) which, as a result, was more likely to win the visual competition. This crossmodal facilitation may be mediated by mid-level lateral excitation, or by top-down feedback connections (Tong et al., 2006). The mid-level lateral excitatory effect can be compared to perceptual grouping (Kovacs et al., 1996; Alais and Blake, 1999; Alais et al., 2006) or the contextual constraints (Treisman, 1962; Shimojo and Nakayama, 1990; Watson et al., 2004) on the perception in the binocular rivalry situation but, in this case, occurring crossmodally. The top-down feedback connection, though, is perhaps the mechanism that auditory semantic context and visual selective attention interactively modulate the visual perception in binocular rivalry.

## HUMAN PERCEPTUAL AWARENESS: UNISENSORY OR MULTISENSORY?

Hupé et al. (2008) recently demonstrated that the perceptual outcomes of simultaneously presented visual and auditory bistable stimuli were generated separately. This result may imply that the sites where conscious perception emerges may be separate for different sensory modalities in terms of the traditional view that each sensory modality has its own processing module (e.g.,

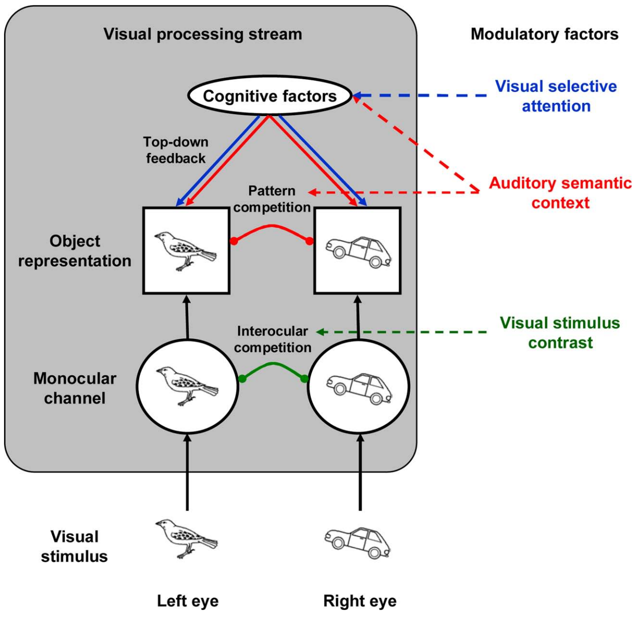

**FIGURE 4 | Schematic figure of the visual processing stream in the binocular rivalry situation adopted from Tong et al. (2006).** In this example, the left eye "sees" a bird while the right eye "sees" a car. The information is represented in monocular- and object-level sequentially, and two types of competition, interocular and pattern competition, occur at each level, respectively. Presumably, the visual stimulus contrast factor modulates interocular competition, whereas the visual selective attention factor modulates pattern competition via top-down feedback connections. The auditory semantic context factor may modulate the pattern competition through either crossmodal excitatory or top-down feedback connections.

Pylyshyn, 1999; Zeki, 2003). Here, on the contrary, we observed that auditory semantic context modulated visual awareness under conditions of binocular rivalry, which is in line with a view of the brain as a closely connected multisensory network: in terms of the neurophysiology, we now know that massive amounts of information is continually being communicated between those brain areas that used to be considered as being sensory-specific. Consequently, many researchers now no longer consider brain regions as being structured as discrete unimodal modules (see Ghazanfar and Schroeder, 2006; Driver and Noesselt, 2008). In terms of psychological functioning, more generally, it is worth considering the powerful constraints that semantics places on the perceptual system as it tries to infer the nature of the environmental stimulation (see Hohwy et al., 2008). That is, audiovisual semantic congruency can provide heuristics, or prior knowledge, on multisensory integration that modulate what we experience on an everyday basis in the real world (i.e., see the literature on the unity assumption, Welch and Warren, 1980; Spence, 2007). The accumulating evidence demonstrating crossmodal semantic interactions in human perception implies that the semantic representations for different sensory modalities are not independent (see McCarthy and Warrington, 1988). However, it is still unclear whether the semantic systems are either completely amodal (e.g., Pylyshyn, 1984), or else the semantic systems for each sensory modality may be highly connected while still retaining some modality-specific information (e.g., Shallice, 1988; Barsalou, 1999; Plaut, 2002). Furthermore, the interplay between perceptual systems and higher-level semantic systems may also imply that perception and cognition share common representation systems, as proposed by the view of grounded cognition (Barsalou, 1999, 2008).

## CONCLUSION

The experiments reported here provide empirical support for the claim that auditory semantic context modulates visual perception in the binocular rivalry situation. The results demonstrate that the effect of auditory semantic context is dissociable from the previously reported effects of visual selective attention and visual stimulus contrast (Meng and Tong, 2004). Recently, the crossmodal modulation of visual perception in the binocular rivalry situation has been demonstrated by the concurrent presentation of both tactile (Lunghi et al., 2010) and olfactory stimuli (Zhou et al., 2010). However, the modulation reported in the former case was based on congruency defined in terms of the direction of motion (see Conrad et al., 2010, for the audiovisual case), while in the latter case it was based on odorant congruency that is comparable to the semantic factor investigated in the present study. We therefore suggest that when considering how the dominant percept in binocular rivalry (and so, human visual awareness) emerges, information from other sensory modalities also needs to be considered; and, in turn, that multisensory stimulation provides a novel means other than unimodal stimulation to probe the contextual constraints on human visual awareness.

## ACKNOWLEDGMENTS

This research was supported by a joint project funded by the British Academy (CQ ROKK0) and the National Science Council in Taiwan (NSC 97-2911-I-002-038). Yi-Chuan Chen was supported by the Ministry of Education in Taiwan (SAS-96109-1-US-37). Su-Ling Yeh receives supports from NSC 96-2413-H-002-009-MY3 and NSC 98-2410-H-002-023-MY3.

## REFERENCES

Alais, D., and Blake, R. (1999). Grouping visual features during binocular rivalry. *Vision Res.* 39, 4341–4353.

Alais, D., and Blake, R. (eds). (2005). *Binocular Rivalry*. Cambridge, MA: MIT Press.

Alais, D., Lorenceau, J., Arrighi, R., and Cass, J. (2006). Contour interactions between pairs of Gabors engaged in binocular rivalry reveal a map of the association field. *Vision Res.* 46, 1473–1487.

Baker, D. H., and Graf, E. W. (2009). On the relation between dichoptic masking and binocular rivalry. *Vision Res.* 49, 451–459.

Balcetis, E., and Dale, R. (2007). Conceptual set as a top-down constraint on visual object identification. *Perception* 36, 581–595.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behav. Brain Sci.* 22, 577–660.

Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645.

Bates, E., D'Amico, S., Jacobsen, T., Székely, A., Andonova, E., Devescovi, A., Herron, D., Lu, C. C., Pechmann, T., Pleh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., Orozco-Figueroa, A., Tzeng, A., and Tzeng, O. (2003). Timed picture naming in seven languages. *Psychon. Bull. Rev.* 10, 344–380.

Biederman, I. (1972). Perceiving real-world scenes. *Science* 177, 77–80.

Blake, R. (1988). Dichoptic reading: the role of meaning in binocular rivalry. *Percept. Psychophys.* 44, 133–141.

Blake, R., and Logothetis, N. K. (2002). Visual competition. *Nat. Rev. Neurosci.* 3, 1–11.

Brand-D'Abrescia, M., and Lavie, N. (2008). Task coordination between and within sensory modalities: effects on distraction. *Percept. Psychophys.* 70, 508–515.

Chen, Y.-C., and Spence, C. (2010). When hearing the bark helps to identify the dog: semantically-congruent sounds modulate the identification of masked pictures. *Cognition* 114, 389–404.

Chen, Y.-C., and Yeh, S.-L. (2008). Visual events modulated by sound in repetition blindness. *Psychon. Bull. Rev.* 15, 404–408.

Chong, S. C., Tadin, D., and Blake, R. (2005). Endogenous attention prolongs dominance durations in binocular rivalry. *J. Vis.* 5, 1004–1012.

Conrad, V., Bartels, A., Kleiner, M., and Noppeney, U. (2010). Audiovisual interactions in binocular rivalry. *J. Vis.* 10, 27.

Costello, P., Jiang, Y., Baartman, B., McGlennen, K., and He, S. (2009). Semantic and subword priming during binocular suppression. *Conscious. Cogn.* 18, 375–382.

Crick, F. (1996). Visual perception: rivalry and consciousness. *Nature* 379, 485–486.

Daselaar, S. M., Porat, Y., Huijbers, W., and Pennartz, C. M. A. (2010). Modality-specific and modality-independent components of the human imagery system. *Neuroimage* 52, 677–685.

Davenport, J. L., and Potter, M. C. (2004). Scene consistency in object and background perception. *Psychol. Sci.* 15, 559–564.

de Gelder, B., and Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends Cogn. Sci. (Regul. Ed.)* 7, 460–467.

Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222.

Driver, J., and Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on "sensory specific" brain regions, neural responses, and judgments. *Neuron* 57, 11–23.

Engel, E. (1956). The role of context in binocular resolution. *Am. J. Psychol.* 69, 87–91.

Ghazanfar, A. A., and Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends Cogn. Sci. (Regul. Ed.)* 10, 278–285.

Glaser, W. R., and Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *J. Exp. Psychol. Gen.* 118, 13–42.

Greene, A. J., Easton, R. D., and LaShell, L. S. R. (2001). Visual-auditory events: cross-modal perceptual priming and recognition memory. *Conscious. Cogn.* 10, 425–435.

Helmholtz, H. V. (1962). *Helmholtz's Treatise on Physiological Optics,* New York, NY: Dover Publications.

Hohwy, J., Roepstorff, A., and Friston, K. (2008). Predictive coding explains binocular rivalry: an epistemological review. *Cognition* 108, 687–701.

Hollingworth, A., and Henderson, J. M. (1998). Does consistent scene context facilitate object perception? *J. Exp. Psychol. Gen.* 127, 398–415.

Hsiao, J. Y. Chen, Y.-C., Spence, C., and Yeh, S.-L. (2010). Semantic congruency, attention, and fixation position modulate conscious perception when viewing a bistable figure. *J. Vis.* 10, 867.

Hupé, J. M., Joffo, L. M., and Pressnitzer, D. (2008). Bistability for audiovisual stimuli: perceptual decision is modality specific. *J. Vis.* 8, 1.1–15.

Iordanescu, L., Guzman-Martinez, E., Grabowecky, M., and Suzuki, S. (2008). Characteristic sound facilitates visual search. *Psychon. Bull. Rev.* 15, 548–554.

Jiang, Y., Costello, P., and He, S. (2007). Processing of invisible stimuli: advantage of upright faces and recognizable words in overcoming interocular suppression. *Psychol. Sci.* 18, 349–355.

Kovacs, I., Papathomas, T. V., Yang, M., and Feher, A. (1996). When the brain changes its mind: interocular grouping during binocular rivalry. *Proc. Natl. Acad. Sci. U.S.A.* 93, 15508–15511.

Leopold, D. A., and Logothetis, N. K. (1996). Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature* 379, 549–553.

Leopold, D. A., and Logothetis, N. K. (1999). Multistable phenomena: changing views in perception. *Trends Cogn. Sci. (Regul. Ed.)* 3, 254–264.

Levelt, W. J. M. (1965). *On Binocular Rivalry*. The Netherlands: Royal VanGorcum.

Logothetis, N. K., Leopold, D. A., and Sheinberg, D. L. (1996). What is rivalling during binocular rivalry? *Nature* 380, 621–624.

Lumer, E. D., Friston, K. J., and Rees, G. (1998). Neural correlates of perceptual rivalry in the human brain. *Science* 280, 1930–1934.

Lunghi, C., Binda, P., and Morrone, M. C. (2010). Touch disambiguates rivalrous perception at early stages of visual analysis. *Curr. Biol.* 20, R143–R144.

McCarthy, R. A., and Warrington, E. K. (1988). Evidence for modality-specific meaning systems in the brain. *Nature* 334, 428–430.

Meng, M., and Tong, F. (2004). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *J. Vis.* 4, 539–551.

Mitchell, J. F., Stoner, G. R., and Reynolds, J. H. (2004). Object-based attention determines dominance in binocular rivalry. *Nature* 429, 410–413.

Molholm, S., Ritter, W., Javitt, D. C., and Foxe, J. J. (2004). Multisensory visual-auditory object recognition in humans: a high-density electrical mapping study. *Cereb. Cortex* 14, 452–465.

Mueller, T. J., and Blake, R. (1989). A fresh look at the temporal dynamics of binocular rivalry. *Biol. Cybern.* 61, 223–232.

Noest, A. J., van Ee, R., Nijs, M. M., and van Wezel, R. J. A. (2007). Percept-choice sequences driven by interrupted ambiguous stimuli: a low-level neural model. *J. Vis.* 7, 10.

Noppeney, U., Josephs, O., Hocking, J., Price, C. J., and Friston, K. J. (2008). The effect of prior visual information on recognition of speech and sounds. *Cereb. Cortex* 18, 598–609.

Ozkan, K., and Braunstein, M. L. (2009). Predominance of ground over ceiling surfaces in binocular rivalry. *Atten. Percept. Psychophys.* 71, 1305–1312.

Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Mem. Cognit.* 3, 519–526.

Pearson, J., Clifford, C. W. G., and Tong, F. (2008). The functional impact of mental imagery on conscious perception. *Curr. Biol.* 18, 982–986.

Plaut, D. C. (2002). Graded modality-specific specialization in semantics: a computational account of optic aphasia. *Cogn. Neuropsychol.* 19, 603–639.

Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behav. Brain Sci.* 22, 341–423.

Pylyshyn, Z. W. (1984). *Computation and Cognition*. Cambridge, MA: MIT Press.

Rommetveit, R., Toch, H., and Svendsen, D. (1968). Effects of contingency and contrast contexts on the cognition of words: a study of stereoscopic rivalry. *Scand. J. Psychol.* 9, 138–144.

Schneider, T. R., Engel, A. K., and Debener, S. (2008). Multisensory identification of natural objects in a two-way crossmodal priming paradigm. *Exp. Psychol.* 55, 121–132.

Schneider, W., Dumais, S. T., and Shiffrin, R. M. (1984). "Automatic and control processing and attention," in *Varieties of Attention*, eds. R. Parasuraman and R. Davies (New York: Academic Press), 1–27.

Segal, S. J., and Fusella, V. (1969). Effects of imaging on signal-to-noise ratio, with varying signal conditions. *Br. J. Psychol.* 60, 459–464.

Segal, S. J., and Fusella, V. (1970). Influence of imagined pictures and sounds on detection of visual and auditory signals. *J. Exp. Psychol.* 83, 458–464.

Shallice, T. (1988). Specialization within the semantic system. *Cogn. Neuropsychol.* 5, 133–142.

Sheth, B. R., and Shimojo, S. (2004). Sound-aided recovery from and persistence against visual filling-in. *Vision Res.* 44, 1907–1917.

Shimojo, S., and Nakayama, K. (1990). Real world occlusion constraints and binocular rivalry. *Vision Res.* 30, 69–80.

Spence, C. (2007). Audiovisual multisensory integration. *Acoust. Sci. Technol.* 28, 61–70.

Spence, C. (2011). Crossmodal correspondences: a tutorial review. *Atten. Percept. Psychophys.* 73, 971–995.

Taylor, K. I., Moss, H. E., Stamatakis, E. A., and Tyler, L. K. (2006). Binding crossmodal object features in perirhinal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 103, 8239–8244.

Tong, F. (2001). Competing theories of binocular rivalry: a possible resolution. *Brain Mind* 2, 55–83.

Tong, F., and Engel, S. A. (2001). Interocular rivalry revealed in the human cortical blind-spot representation. *Nature* 411, 195–199.

Tong, F., Meng, M., and Blake, R. (2006). Neural bases of binocular rivalry. *Trends Cogn. Sci. (Regul. Ed.)* 10, 502–511.

Tong, F., Nakayama, K., Vaughan, J. T., and Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron* 21, 753–759.

Treisman, A. (1962). Binocular rivalry and stereoscopic depth perception. *Q. J. Exp. Psychol.* 14, 23–37.

Treisman, A., and Davies, A. (1973). "Dividing attention to ear and eye," in *Attention and Performance IV*, ed. S. Kornblum (New York: Academic Press), 101–117.

Treisman, A., and Riley, J. G. A. (1969). Is selective attention selective perception or selective response? A further test. *J. Exp. Psychol.* 79, 27–34.

van Atteveldt, N., Formisano, E., Goebel, R., and Blomer, L. (2004). Integration of letters and speech sounds in the human brain. *Neuron* 43, 271–282.

van Boxtel, J. J. A., van Ee, R., and Erkelens, C. J. (2007). Dichoptic masking and binocular rivalry share common perceptual dynamics. *J. Vis.* 7, 3.1–11.

van Ee, R., van Boxtel, J. J. A., Parker, A. L., and Alais, D. (2009). Multimodal congruency as a mechanism for willful control over perceptual awareness. *J. Neurosci.* 29, 11641–11649.

van Ee, R., van Dam, L. C. J., and Brouwer, G. J. (2005). Voluntary control and the dynamics of perceptual bi-stability. *Vision Res.* 45, 41–55.

Watson, T. L., Pearson, J., and Clifford, C. W. G. (2004). Perceptual grouping of biological motion promotes binocular rivalry. *Curr. Biol.* 14, 1670–1674.

Welch, R. B., and Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychol. Bull.* 3, 638–667.

Yang, Y. H., and Yeh, S.-L. (2011). Accessing the meaning of invisible words. *Conscious. Cogn.* 20, 223–233.

Yuval-Greenberg, S., and Deouell, L. Y. (2009). The dog's meow: asymmetrical interaction in cross-modal object recognition. *Exp. Brain Res.* 193, 603–614.

Zeki, S. (2003). The disunity of consciousness. *Trends Cogn. Sci. (Regul. Ed.)* 7, 214–218.

Zhou, W., Jiang, Y., He, S., and Chen, D. (2010). Olfaction modulates visual perception in binocular rivalry. *Curr. Biol.* 20, 1356–1358.

Zimba, L. D., and Blake, R. (1983). Binocular rivalry and semantic processing: out of sight, out of mind. *J. Exp. Psychol. Hum. Percept. Perform.* 9, 807–815.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 17 June 2011; accepted: 15 August 2011; published online: 12 September 2011.*

*Citation: Chen Y-C, Yeh S-L and Spence C (2011) Crossmodal constraints on human perceptual awareness: auditory semantic modulation of binocular rivalry. Front. Psychology* ***2****:212. doi: 10.3389/fpsyg.2011.00212*

*This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.*

frontiers in
PSYCHOLOGY

**PERSPECTIVE ARTICLE**
published: 18 October 2011
doi: 10.3389/fpsyg.2011.00264

# Influences of multisensory experience on subsequent unisensory processing

***Ladan Shams*[1,2]*, *David R. Wozny*[2,3], *Robyn Kim*[1] *and Aaron Seitz*[4]**

[1] *Department of Psychology, University of California Los Angeles, Los Angeles, CA, USA*
[2] *Biomedical Engineering Interdepartmental Program, University of California Los Angeles, Los Angeles, CA, USA*
[3] *Department of Otolaryngology, Oregon Health and Science University, Portland, OR, USA*
[4] *Department of Psychology, University of California Riverside, Riverside, CA, USA*

***Edited by:***
*Nadia Bolognini, University Milano-Bicocca, Italy*

***Reviewed by:***
*Thomas J. Perrault Jr., Wake Forest University, USA*
*Francesco Marini, University of Milano-Bicocca, Italy*

***Correspondence:***
*Ladan Shams, Department of Psychology, University of California Los Angeles, Los Angeles, CA 90095-1563, USA.*
*e-mail: ladan@psych.ucla.edu*

Multisensory perception has been the focus of intense investigation in recent years. It is now well-established that crossmodal interactions are ubiquitous in perceptual processing and endow the system with improved precision, accuracy, processing speed, etc. While these findings have shed much light on principles and mechanisms of perception, ultimately it is not very surprising that multiple sources of information provides benefits in performance compared to a single source of information. Here, we argue that the more surprising recent findings are those showing that multisensory experience also influences the subsequent *unisensory* processing. For example, exposure to auditory–visual stimuli can change the way that auditory or visual stimuli are processed subsequently even in isolation. We review three sets of findings that represent three different types of learning ranging from perceptual learning, to sensory recalibration, to associative learning. In all these cases exposure to multisensory stimuli profoundly influences the subsequent unisensory processing. This diversity of phenomena may suggest that continuous modification of unisensory representations by multisensory relationships may be a general learning strategy employed by the brain.

**Keywords: multisensory integration, multisensory representation, unisensory representation, multisensory learning, learning facilitation**

## INTRODUCTION

We live in a world that is replete with multisensory information. As such, multisensory processing has been an active topic of research and numerous studies have demonstrated that multisensory processing can improve accuracy (e.g., Sumby and Pollack, 1954, reduce reaction times, e.g., Gingras et al., 2009), improve precision (e.g., Ernst and Banks, 2002; Alais and Burr, 2004), and provide more complete information about objects (Newell et al., 2001). Furthermore, recent studies have established the presence of a significant degree of plasticity in multisensory processes, including processes such as crossmodal simultaneity (e.g., Fujisaki et al., 2004, and temporal order, e.g., Miyazaki et al., 2006) that had previously been thought to be hardwired or highly stable. However, how multisensory processing impacts subsequent *unisensory* processing has received less attention. This is despite the fact that several studies indicate that unisensory processing is altered through multisensory experience.

In Section "Improvement in Unisensory Sensitivity as a Result of Correlated Multisensory Training," we describe recent studies that show that training observers using correlated auditory–visual stimuli improves subsequent performance in a unisensory (visual or auditory) detection, discrimination, and recognition task. In Section "Change in Unisensory Map as a Result of Exposure to Crossmodal Error," we discuss recent research demonstrating that momentary exposure to auditory–visual spatial discrepancy results in a shift in the auditory space map. We discuss how this crossmodal sensory recalibration is continuously engaged in updating unisensory perceptual processing and is an integral part of perceptual processing. In Section "Improvement in Unisensory Sensitivity as a Result of Multisensory Associative Learning," we present results from an adaptation study that shows that passive exposure to consistently paired auditory and visual features enhances visual sensitivity. These three sets of findings involve very different types of learning – perceptual learning, recalibration, and associative learning – and may involve different mechanisms and time scales, yet they all show a significant influence of multisensory processing on unisensory representations. This diversity of phenomena suggests that these multisensory influences on unisensory learning may reflect a general strategy of learning in the brain.

## IMPROVEMENT IN UNISENSORY SENSITIVITY AS A RESULT OF CORRELATED MULTISENSORY TRAINING

Multisensory stimulation is widely thought to be advantageous for learning (Montessori, 1912; Fernald and Keller, 1921; Orton, 1928; Strauss and Lehtinen, 1947). As such, numerous educational programs, including the Montessori (1912, 1967) and Multisensory Structural Language Education method (Birsh, 1999), incorporate multisensory training techniques in their teaching. The benefits of multisensory training go beyond the simultaneous engagement of individuals with different learning styles (e.g., "visual learners" and "auditory learners"; Coffield et al., 2004). However, benefits of multisensory training are typically stated in anecdotal terms, such as Treichler's (1967) statement that "People generally remember 10% of what they read, 20% of what they hear, 30% of what they see, and 50% of what they see and hear." (but see Thompson and Paivio, 1994). While the benefits of multisensory training have long been appreciated and exploited by educational and clinical practitioners, until recently there has been little solid scientific evidence to support this view.

To address the extent to which multisensory training shows benefits over unisensory training, we recently investigated how visual perceptual learning of motion–direction perception (Ball and Sekuler, 1982, 1987; Liu, 1999; Seitz et al., 2006a,b; Chalk et al., 2010; Pilly et al., 2010) is influenced by the addition of auditory information (Seitz et al., 2006a; Kim et al., 2008). Perceptual learning is an appropriate method to address the benefits of multisensory training since it is a well-established learning paradigm and a great deal is known regarding the mechanisms involved (Gilbert et al., 2001; Fahle and Poggio, 2002; Ahissar and Hochstein, 2004; Ghose, 2004; Seitz and Dinse, 2007; Shams and Seitz, 2008). We compared the effects of congruent auditory–visual (AVcong-trained) and visual (V-trained) training on perceptual learning using a coherent motion detection and discrimination task (Seitz et al., 2006a). The individuals in the AVcong-trained group were trained using auditory and visual stimuli moving in the same direction, where as the V-trained group was trained only with visual motion stimuli. Critically, the two groups were compared on trials *without* informative auditory signals (stationary sound, and in a subsequent study described below, the two groups were compared on identical trials with no sound). Compared to the V-trained group, the AVcong-trained group showed greater learning both within the first session and across the 10 training sessions (**Figure 1A**). Therefore, *multisensory training facilitated unisensory learning*. The advantage of AV training over visual-alone training was substantial: it reduced the number of sessions required to reach asymptote by ~60%, while also raising the maximum performance.

A second study (Kim et al., 2008) showed that benefits of multisensory training were specific to training with congruent auditory–visual stimuli (i.e., moving in the same direction); a group trained with sound moving in the opposite direction of visual motion (AVincong-trained group) did not show any facilitation of learning (**Figure 1B**). This indicates that the facilitation of learning is not due to a putative alerting effect of sound during training. Additionally, results of a direction test showed that performance was significantly greater for trained directions (10° and 190°) than for untrained directions, confirming that this improvement reflects perceptual learning rather than general task learning (Ball and Sekuler, 1982; Fahle, 2004). Intriguingly, for the AVcong-trained group, the performance on silent visual trials (**Figure 1B**, solid blue) converged to the level of performance on congruent AV trials (**Figure 1B**, broken blue). In other words, individuals trained with congruent AV stimuli not only showed facilitated visual performance *when auditory stimuli were not present*, but also they performed in the absence of sound *as well as* they would perform in the presence of sound.

Other studies demonstrate that these beneficial effects are not limited to visual perceptual learning. For example, individuals trained with faces and voices can better recognize voices (auditory-alone) than those trained with voices alone (Von Kriegstein and Giraud, 2006). Memory research suggests that multisensory encoding of objects facilitates the subsequent retrieval of unisensory information (Murray et al., 2004, 2005; Lehmann and Murray, 2005). In addition, multisensory exposure has been reported to enhance unisensory reinforcement learning (Guo and Guo, 2005) in Drosophila (fruit flies). Collectively these studies indicate that crossmodal facilitation of learning is a general phenomenon occurring in different tasks, and across different modalities, and even species.

In a recent review, Shams and Seitz (2008) discussed how multisensory training could benefit later performance of unisensory tasks. It was suggested that facilitation could arise through two

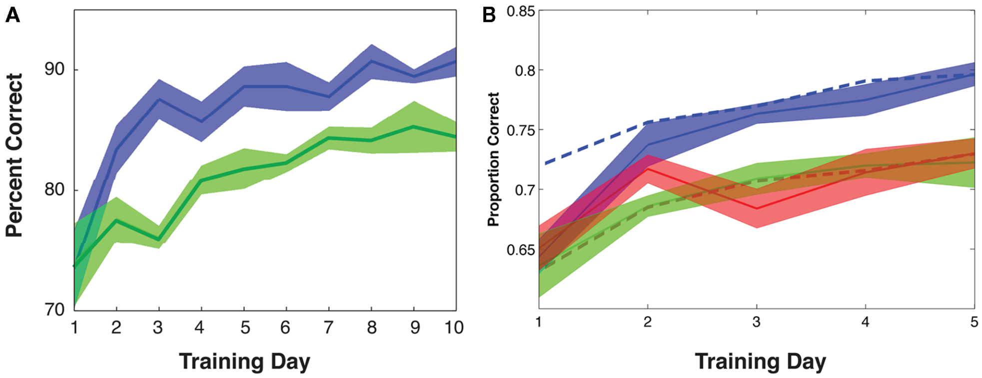

**FIGURE 1 | Benefits of multisensory training to visual learning.** Performance on visual-only trials (no auditory signal) is shown for different groups trained in different conditions. Green, blue, and red curves represent groups trained with only visual stimuli (V-trained group), trained with congruent auditory and visual motion (AVcong-trained group), and trained with incongruent (moving in opposite directions) auditory and visual motion (AVincong-trained group), respectively. **(A)** Data from Seitz et al. (2006a) shows that learning occurred more quickly and more extensively for the AVcong-trained group (blue) compared to the V-trained group (green). Figure adapted from Seitz et al. (2006a) with permission. **(B)** Data from Kim et al. (2008) shows that relative to the V-trained group (green), the enhanced learning is limited to the AVcong-trained group (solid blue) and does not occur for the AVincong-trained group (solid red). Solid lines represent performance on silent V trials. Broken blue and broken red lines show performance in congruent AV and incongruent AV trials, respectively. Figure adapted from Kim et al. (2008).

classes of mechanisms. One possibility is that facilitation benefits learning in the same representations that undergo modification in classic unisensory learning (Seitz and Dinse, 2007). Alternatively, facilitation can be explained through multisensory exposure resulting in alterations to multisensory representations that can then be invoked by a unisensory component (Rao and Ballard, 1999; Friston, 2005). While, the findings discussed in this section can be explained by either, or a combination, of these mechanisms, other findings discussed below are suggestive that the latter mechanism (unisensory representations becoming equivalent to multisensory representations) likely play some role in the observed facilitation of learning.

## CHANGE IN UNISENSORY MAP AS A RESULT OF EXPOSURE TO CROSSMODAL ERROR

As highlighted in the introduction, being endowed with multiple sensory modalities has its advantages in immediate perceptual processing. However, as illustrated in the previous section, multisensory stimulation also has a lasting effect on subsequent unisensory stimulation. This section describes the phenomenon of crossmodal sensory recalibration. Perception can generally be considered an unsupervised inference process, where the ground truth (i.e., the environmental state) is unknown, and can only be estimated from the sensorium. Therefore, comparing sensory estimates across modalities over time allows the system to perform self-maintenance by recalibrating its unisensory processes (King, 2009; Recanzone, 2009). Such changes are necessary when coping with endogenous changes that occur during development or injury, or exogenous changes in environmental conditions. An example of crossmodal recalibration is the rubber-hand illusion in which a brief (seconds) tactile stimulation of one's occluded arm while seeing a synchronous tactile stimulation of a rubber-hand subsequently induces a shift in the proprioception of the hand in the direction of the seen rubber hand (Botvinich and Cohen, 1998). Another extensively studied example of crossmodal recalibration is the ventriloquist aftereffect (VAE): the shift in perceived location of sounds (in isolation) that occurs after repeated exposure to consistent spatial discrepancy between auditory and visual stimuli (Canon, 1970; Radeau and Bertelson, 1974; Recanzone, 1998; Lewald, 2002).

While the rubber-hand illusion shows that recalibration of proprioception can occur rapidly, after seconds of exposure to tactile–visual discrepancy, recalibration of other sensory modalities such as hearing and vision has been shown to occur only after substantial exposure to spatial inconsistencies between the sensory signals, for example, after hundreds or thousands of repeated exposures to consistent discrepancy between the senses (Radeau and Bertelson, 1974; Zwiers et al., 2003; Navarra et al., 2009). In some cases, auditory recalibration has been reported after weeks, days, or hours of exposure to inconsistency (Hofma et al., 1998; Zwiers et al., 2003). The VAE has been reported to occur after several minutes of continuous exposure, or after thousands or hundreds of trials (Canon, 1970; Radeau and Bertelson, 1974; Recanzone, 1998; Lewald, 2002; Frissen et al., 2003). Altogether these results have given the impression that the human auditory and visual systems require a substantial amount of evidence that the sense is faulty before recalibration occurs.

Wozny and Shams (2011) recently conducted a study that demonstrated that auditory–visual spatial recalibration occurs much more quickly than previously thought. Observers were presented with small white disks on a black screen and white noise bursts at variable locations along azimuth for 35 ms, and were asked to localize the stimuli using a trackball that controlled the position of a cursor on the screen. On some trials only an auditory stimulus was presented, on some trials only a visual stimulus was presented, and on some trials both were presented. On bisensory trials, the observers were asked to report the location of both the visual stimulus and the auditory stimulus. All combinations of visual and auditory visual locations were presented with equal probability on both unisensory and bisensory trials, and the trials were interleaved pseudorandomly. Therefore, an auditory-alone trial could be preceded by a visual, auditory, or auditory–visual trial, and the spatial discrepancy between the auditory and visual stimuli could vary from trial to trial. This experimental design allowed us to investigate whether there is a systematic influence of AV spatial discrepancy experienced on a bisensory trial on the subsequent perception of location of sound on a unisensory auditory trial. In **Figure 2**, the change in perceived location of sound is plotted as a function of AV discrepancy in the immediately preceding AV trial. As can be seen, the perceived location of sound is shifted to the right if the auditory trial is preceded by a trial in which vision is to the right of sound, and the perceived location of sound is shifted to the left if the auditory trials is preceded by a trial in which visual stimulus was to the left of the auditory stimulus. The shift in perceived location is calculated as a difference between the reported location on a given auditory trial as compared to the reported location of sound averaged across all unisensory auditory trials with sound presented at the same location. The same qualitative results are obtained if change in perceived location is measured relative to the actual location of sound.

These findings show that auditory recalibration can occur very rapidly, after only *milliseconds* of exposure to sensory discrepancy and suggest that any exposure to discrepant auditory–visual

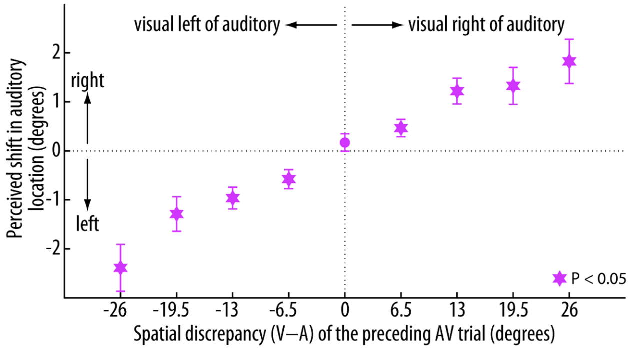

**FIGURE 2 | Shift in auditory map as a function of specific exposures in the preceding trial.** The shift in perceived auditory location (mean ± SEM across observers) as a function of auditory–visual spatial discrepancy in the preceding AV trial. Stars denoted datapoints that are significantly different from zero (corrected for multiple comparisons using Bonferroni–Holm correction). Figure reproduced from Wozny and Shams (2011) with permission.

sensations can instantaneously change the subsequent perception of location of sounds. This indicates a much stronger degree of malleability in our basic auditory representations (such as space) than previously thought.

Interestingly, the degree of recalibration appears to depend more on the perceived discrepancy between the auditory and visual stimuli than the physical discrepancy. The amount of recalibration was four times larger for trials in which the auditory and visual stimuli were perceived to originate from the same location than in trials where they appeared to stem from different locations (Wozny and Shams, 2011). Considering that it is not clear how long lasting the observed shifts in the auditory map are, it is possible that the recalibration phenomenon discussed here and the learning effects discussed in the previous section are mediated by distinct neural mechanisms. Studies in barn owls have found that audio–visual recalibration can involve plasticity in traditionally considered unisensory auditory and visual brain areas such as inferior colliculus (Feldman and Knudsen, 1997) and optic tectum (DeBello and Knudsen, 2004). Whether the rapid human spatial recalibration observed in Wozny and Shams (2011) involves similar mechanisms is a target of future research.

## IMPROVEMENT IN UNISENSORY SENSITIVITY AS A RESULT OF MULTISENSORY ASSOCIATIVE LEARNING

While the studies described above detail how unisensory representations are altered through multisensory experience, they do not directly address how the unisensory processing is impacted by the presence of the multisensory stimulation. In a recent study, Wozny et al. (2008) investigated[1] whether after exposure to arbitrarily paired auditory and visual features, the processing of the visual feature is enhanced by the mere accompaniment of the associated auditory feature even when auditory signals are not informative for the task. If the learning of auditory–visual associations occurs at a sensory level, one could expect that the mere presence of the associated auditory feature could improve the representation of the visual feature, however if the association is not established or if it is established at a higher level of processing, then the presence of task-irrelevant auditory signal would not enhance the visual performance (detection, discrimination, etc.). To address this issue, two experiments were conducted in which observers were passively exposed to a paired auditory–visual stimulus. In both experiments, observers demonstrated a relative increase in sensitivity to that visual stimulus when it was accompanied by the auditory stimulus that was coupled with it during exposure, even though auditory stimulus was uninformative to the subjects' task. These results suggest that unisensory benefits occur, at least in part, due to an alteration, or formation, of multisensory representations of the stimuli, as discussed in Shams and Seitz (2008).

In one experiment, oriented sinusoidal gratings were paired with pure tones. During the exposure phase, a sinusoidal grating of given visual angle of orientation (V1) was consistently presented with an auditory tone (A1) while the orthogonal orientation (V2) was presented in silence (**Figure 3A**). The visual and auditory stimuli (V1A1) co-varied in randomly chosen suprathreshold stimulus intensities across trials. The task was to keep fixation and detect any changes in the color of the fixation cross by pressing the spacebar. A change in fixation cross color occurred in approximately 10% of trials. Testing occurred prior to and after exposure. During test sessions subjects had to detect in which of two intervals the oriented grating appeared (embedded in visual noise). In trial types that involved the presentation of tones, the tone was played in both intervals and therefore, was uninformative for the task. Each test session consisted of 192 randomly interleaved trials (48 per condition). Subjects who scored close to chance (below 60%) on one or more of the pre-test conditions were excluded from sample. The two exposure conditions and four testing conditions are shown in **Figure 3A**.

A two-way repeated measures ANOVA with factors Test (pre and post) and Condition (V1A1, V1A2, V1, and V2A1) showed a significant interaction between Test and Condition [$F(3,114) = 3.86$, $p < 0.05$]. To determine whether passive exposure to a specific pair of auditory and visual stimuli would result in a relative increase in detection performance for that visual

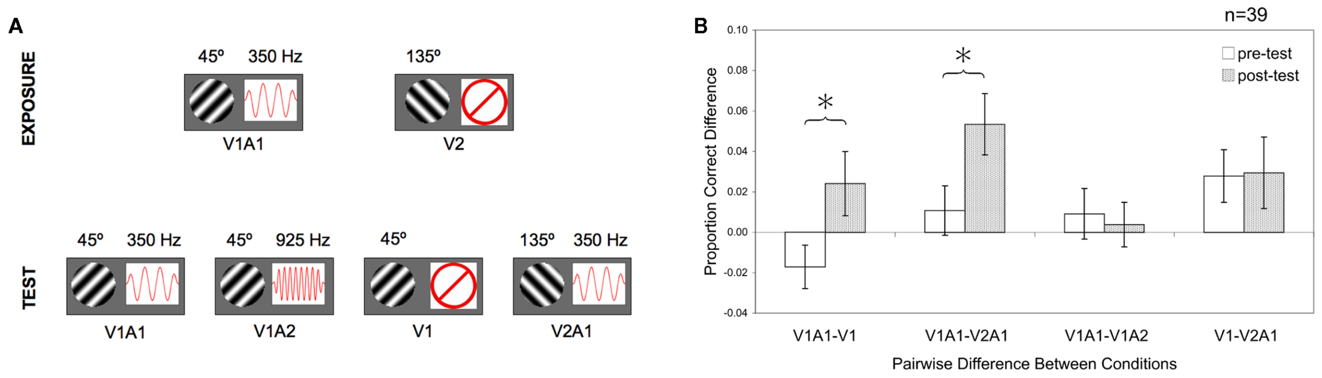

**FIGURE 3 | Influence of exposure to paired visual orientation and auditory frequency on subsequent visual orientation detection. (A)** Top, The stimulus conditions to which the subjects were passively exposed. Bottom, the stimulus conditions in which subjects were tested in a 2IFC detection task. **(B)** The difference in performance between conditions before and after exposure. Stars denote significant one-tailed paired *t*-tests ($p < 0.05$) between pre and post tests corrected for multiple comparisons using the Bonferroni–Holm method. Error bars represent SE.

---

[1] These results were presented at the 2008 Vision Sciences Society Meeting and an abstract of the study is published in Journal of Vision as cited in the text.

stimulus when accompanied by the associated sound, we compared performance differences between the pre-test and post-test data between V1A1 and V1, and found that there was a significant difference between these conditions ($p = 0.013$, one-tailed paired $t$-test, df = 38, Bonferroni–Holm $\alpha = 0.017$; **Figure 3B** column 1). If the pairing with sound had only facilitated the visual learning, the relative performance between these two conditions should have been the same. In contrast, our results suggest that an auditory–visual association was learned.

To determine whether the benefit for the V1A1 condition is a specific effect to this associated auditory–visual stimulus or whether it is a generalized effect, we examined the performance on the other testing conditions. First, if the improved performance in V1A1 is due to an alerting effect of sound, then we would expect to see the same degree of improvement in both V1A1 and V2A1. However, this was not the case as the comparison between V1A1 vs. V2A1 conditions confirmed that the facilitation was orientation specific ($p = 0.009$, one-tailed paired $t$-test, df = 38, Bonferroni–Holm $\alpha = 0.0125$; **Figure 3B** column 2). However, a significant difference was not found between learning for the exposed V1A1 condition (350 Hz tone) vs. the same orientation paired with a slightly different tone V1A2 (925 Hz), suggesting that the learning transfers across at least some range of frequencies (**Figure 3B** column 3). This degree of transfer is not entirely surprising given that the frequencies of A1 and A2 lie within an octave and a half of each other, which is within the range of auditory recalibration transfer shown in other studies (Frissen et al., 2003). Future experiments should investigate whether a wider frequency range would still show transfer of learning. Finally, as a control, we compared two conditions that had an equal amount of exposure to their components, but arranged in opponent pairings (V1 vs. V2A1) and found there was no noticeable difference in relative performance across these conditions (**Figure 3B** column 4). Altogether, these results suggest that a specific auditory–visual association was learned between V1 and A1 by passive exposure.

In the experiment described above, the auditory–visual pairing presented to subjects during exposure (V1A1) showed the greatest degree of relative improvement. This condition also happened to be the only condition tested in which the visual stimulus was presented in the same context as that of the exposure phase. Therefore a similarity in context can be an alternative explanation for the pattern of results found in the first experiment. To address this potential confound, and to see if the effect can be replicated with other visual features, we conducted a second experiment. In this experiment, the oriented gratings were replaced by coherent dot motion. The exposure phase was similar to the first experiment, where an auditory tone (A1) was consistently paired with a particular direction of coherent motion (V1), while the orthogonal motion–direction (V2) was presented in silence. During testing, subjects had to determine the direction of coherent motion, presented with and without A1. Schematic depiction of the design is shown in **Figure 4A**, which shows the testing and exposure pairings. In contrast to the first experiment, here in addition to testing the exposed auditory visual pair V1A1, we tested V2, in which the other visual feature (not coupled with sound) is also presented in the same context (no sound) as that of the exposure phase. If the improved performance in V1A1 observed in the first experiment was due to familiar context, then similar improvement should be observed here for V2 (no-sound context). But if the improved performance was due to acquisition of a compound AV feature, then the improvement should only be observed for V1A1 and not for V2.

The exposure phase was very similar to that of Experiment 1. Subjects were presented with two trial types: V1A1 and V2. Four hundred trials of each condition were presented in pseudorandom order. Subjects were instructed to maintain fixation and to report any changes in the contrast of the fixation dot. Exposure was preceded by 256 test trials, and followed by 128 randomly interleaved test trials, 400 more exposure trials, and 128 test trials. This top-up design was used to minimize the erosion of learning effect during post-test trials. The post-test results shown below reflect the data from all 256 post-exposure trials. The entire experiment lasted about an hour. For the test sessions, a two-alternative-forced-choice (2AFC) procedure was used where a single trial was presented and the subjects were asked to report by keypress whether the coherent motion moved at 45° or 135°. Four stimulus

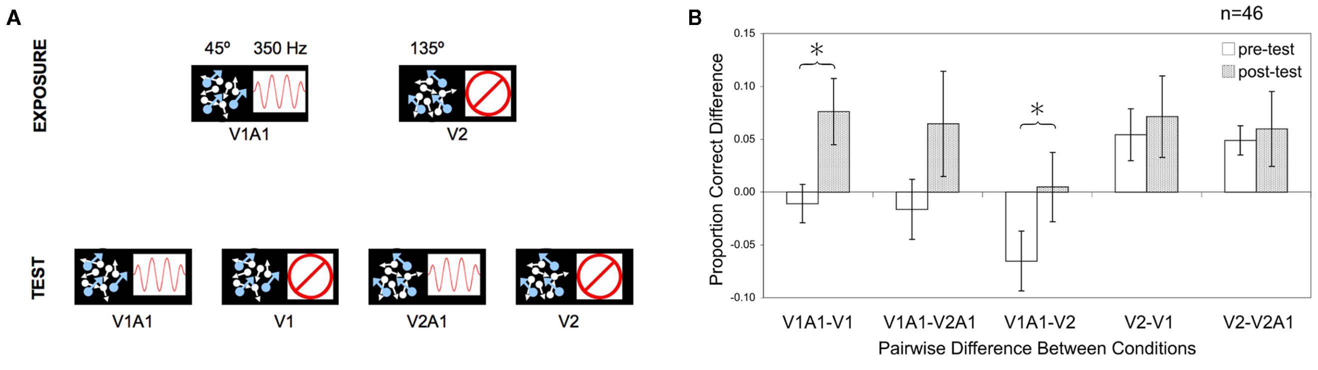

**FIGURE 4 | Influence of exposure to paired visual motion–direction and auditory frequency on subsequent visual motion detection. (A)** Top, The stimulus conditions to which the subjects were passively exposed. Bottom, the stimulus conditions in which subjects were tested in a 2IFC detection task. **(B)** The difference in performance between conditions before and after exposure. Stars denote significant one-tailed paired $t$-tests ($p < 0.05$) between pre and post tests corrected for multiple comparisons using the Bonferroni–Holm method. Error bars represent SE.

conditions were tested: V1A1, V1, V2A1, and V2. Therefore, sound was not informative for the task.

This experiment replicated the findings of the first experiment. Similar to the previous experiment, we performed two-way repeated measures ANOVA with Test (pre, post) and Condition (V1A1, V1, V2A1, V2) as factors. We found a significant interaction $[F(3,135) = 2.68, p < 0.05]$. Here too, there was a significant difference between conditions V1A1 and V1 ($p = 0.007$, one-tail paired $t$-test, df = 45, Bonferroni–Holm $\alpha = 0.01$; **Figure 4B** column 1). This effect seems to be direction specific given that there is a trend of increased performance in the V1A1 conditions compared to V2A1 condition ($p = 0.036$, one-tailed paired $t$-test, df = 45, Bonferroni–Holm $\alpha = 0.0167$; **Figure 4B** column 2). The fact that the results hold true for a discrimination task in addition to the detection task used in the first experiment demonstrates that these effects are not a task-specific oddity. The fact that the V1A1 association is found for motion–direction stimuli in addition to static oriented gratings suggests that these automatic associations that we observe between the auditory and visual stimuli are a general visual phenomenon.

Another goal of this experiment was to test the hypothesis that the learning was simply due to shared context with the exposure, rather than an effect that depended on multisensory stimulation. To address this question we compared the performance between the two tested contexts that were maintained from the exposure (i.e., V1A1 and V2). We found that performance improvement from pre-test to post-test in V1A1 was superior compared to performance improvement in V2 ($p = 0.012$, one-tail paired $t$-test, df = 45, Bonferroni–Holm $\alpha = 0.0125$; **Figure 4B**, column 3). Likewise, columns 4 and 5 of **Figure 4B** show comparison V2 vs. V1 and V2 vs. V2A1, respectively. There was not any significant difference between these conditions, even though the V2 condition was equally exposed as the V1A1 condition. These results confirm that the presentation in familiar context is not the underlying factor behind the observed improvements for V1A1.

A key question is whether the exposure period creates a response bias or leads to a change in sensitivity to the stimulus. In the first experiment, we used a 2IFC paradigm in which response bias has no impact on the results. In the second experiment, we found an increase in sensitivity for the AV trials after exposure (**Figure 5A**) and no change in the bias measurements (**Figure 5B**). Our 2IFC design for the first experiment and signal detection analysis for the second experiment indicate that the improved relative performance observed for the detection/discrimination of the sound-coupled visual feature is due to an increase in sensitivity. This finding in turn suggests that the improved performance reflects learning of a low-level perceptual association. These results therefore suggest that new auditory–visual perceptual associations can be acquired based on brief exposure to correlated auditory and visual coincidences even in adult sensory systems. This indicates an impressive degree of plasticity across modalities in early sensory processing.

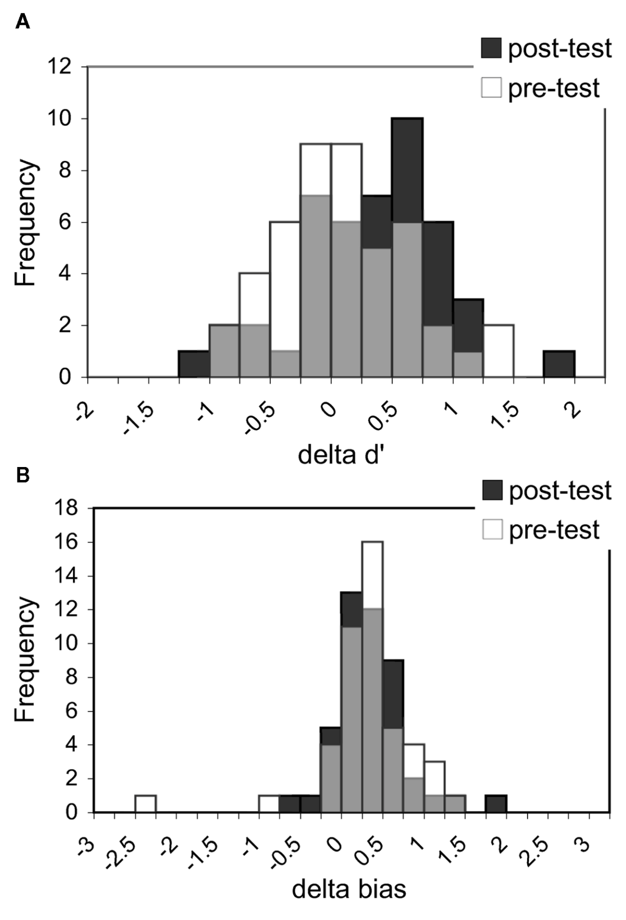

**FIGURE 5 | Signal detection analysis of the experiment on associative learning of visual motion and auditory frequency. (A)** A histogram of $d'$ differences between the auditory–visual (AV) and the vision alone conditions (V). White and dark bars show the pre- and post-test AV–V $d'$ frequencies, respectively. Light gray bars show an overlap in distributions. For each subject, the $d'$ values of the vision alone trials were subtracted from that of the auditory–visual trials. Thus, a positive shift in the distribution indicates an increase in sensitivity for the AV trials. **(B)** Histogram showing frequencies of bias measurements for the AV–V conditions.

In contrast to previous studies of crossmodal associative learning, our study compares the effect of crossmodal associative learning on sensitivity to a visual feature with that of an exposure to the visual stimulus alone. The fact that improvement in V1A1 condition was superior to that of V2 – despite the equal exposure of V1 and V2 – indicates that the increase in sensitivity to a visual feature achieved through establishment of a new auditory–visual feature is superior to any fine tuning of the representation obtained by exposure to the visual feature alone. This is an interesting finding, and can have important implications for perceptual skill acquisition in general. The exact mechanism by which the coupling of sound with the visual stimulus results in improved detection and discrimination of the visual stimuli is not clear. However, one possible mechanism is one in which the correlated incidence of the auditory and visual stimuli leads to establishment of new connections between the two types of feature detectors, i.e., the formation of a multisensory representation (Shams and Seitz, 2008). This will result in increased gain in the visual feature detectors whenever the visual stimulus is encountered in presence of the coupled sound. The increase in gain will in turn result in a higher sensitivity to the visual stimulus. Future studies will need to test this hypothesis.

## DISCUSSION AND CONCLUSION

The human brain has evolved to learn and operate optimally in natural environments in which behavior is guided by information integrated across multiple sensory modalities. Crossmodal interactions are ubiquitous in the nervous system and occur even at early stages of perceptual processing (Shimojo and Shams, 2001; Calvert et al., 2004; Schroeder and Foxe, 2005; Ghazanfar and Schroeder, 2006; Driver and Noesselt, 2008). Until recently, however, studies of perceptual learning focused on training with one sensory modality. This unisensory training fails to tap into natural learning mechanisms that have evolved to optimize behavior in a multisensory environment.

We discussed three sets of learning phenomena that differ both in time scale and type of learning. However, in all cases multisensory exposure caused a marked change in later unisensory processing. In the learning studies discussed in Section "Improvement in Unisensory Sensitivity as a Result of Correlated Multisensory Training," the facilitation of visual learning by sound was apparent within the first hour-long session as well as across days of training. In the experiments discussed in Section "Improvement in Unisensory Sensitivity as a Result of Multisensory Associative Learning," the visual learning was evident after minutes of exposure to paired auditory–visual stimuli. The crossmodal recalibration study discussed in Section "Change in Unisensory Map as a Result of Exposure to Crossmodal Error" provided evidence that significant changes in unisensory representations can occur after only milliseconds of exposure to conflicting auditory–visual stimuli.

In the recalibration study discussed in Section "Change in Unisensory Map as a Result of Exposure to Crossmodal Error," as well as many other previous studies of crossmodal recalibration, a mismatch between two sensory modalities (or in sensorimotor modalities) causes a change in unisensory representations. The study by Wozny and Shams (2011) shows that this adjustment of unisensory representation based on an error signal computed from comparison with another modality does not require a protracted exposure to repeated error, and occurs continuously and incrementally. This continuous modification of unisensory representations as a result of exposure to crossmodal mismatch blurs the distinction between unisensory processing and multisensory processing. It appears that unisensory representations are closely yoked to mechanisms that keep track of crossmodal consistency/error even in the mature human nervous system.

In contrast to the learning involved in recalibration, which is caused by exposure to a mismatch between modalities, the learning phenomena discussed in Sections "Improvement in Unisensory Sensitivity as a Result of Correlated Multisensory Training" and "Improvement in Unisensory Sensitivity as a Result of Multisensory Associative Learning" result from exposure to multisensory stimuli that are not mismatched. In both of these cases, exposure to correlated auditory–visual stimuli causes enhanced performance in unisensory tasks. In the perceptual learning studies discussed in Section "Improvement in Unisensory Sensitivity as a Result of Correlated Multisensory Training," the multisensory stimuli are ecologically correlated, whereas in the associative learning experiments of Section "Improvement in Unisensory Sensitivity as a Result of Multisensory Associative Learning" the pairing between the stimuli is arbitrary (see also Ernst, 2007). We suggest that associative and perceptual learning may represent two different stages of learning along the same dimension, with the associative learning (see Improvement in Unisensory Sensitivity as a Result of Multisensory Associative Learning) representing an initial process of learning and the perceptual learning (see Improvement in Unisensory Sensitivity as a Result of Correlated Multisensory Training) occurring once the association is built (see **Figure 6**). The idea is that initially the auditory and visual stimuli are not associated with each other in the brain,

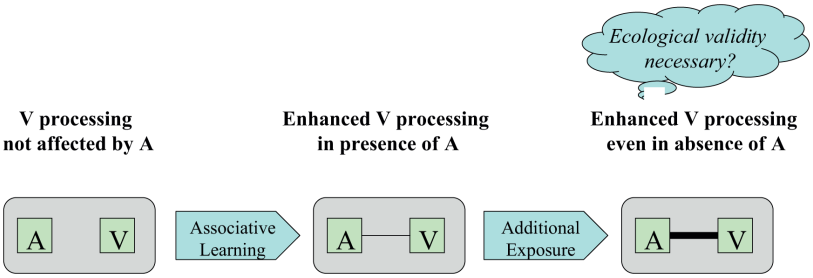

**FIGURE 6 | A possible progression of learning as a result of repeated exposure to coupled auditory and visual stimuli.** The representation of auditory and visual stimuli are initially not linked in the brain. Repeated exposure to paired auditory and visual stimuli results in associative learning. The newly learned association between the auditory and visual features (A and V) results in enhanced processing of the visual stimuli when accompanied by the coupled auditory stimuli. This phenomenon was discussed in Section "Improvement in Unisensory Sensitivity as a Result of Multisensory Associative Learning." For auditory and visual stimuli that are already associated in the brain, additional repeated exposure causes the connectivity/association between the two features to be strengthened further, gradually blurring the distinction between unisensory and bisensory representations (a unisensory representation becomes as effective as a bisensory representation). This strong link between the two representations results in enhanced processing of the visual features even in the absence of the coupled auditory stimulation (and vice versa). This phenomenon was discussed in Section "Improvement in Unisensory Sensitivity as a Result of Correlated Multisensory Training." However, alternatively, the learning of association between arbitrary A and V stimuli may not progress to the phenomenon of enhanced visual processing in the absence of A. The latter phenomenon may be confined to A and V features that are ecologically related (such as motion) as it may require hard-wiring between brain areas that mediate their representations. If so, the phenomenon discussed in Sections "Improvement in Unisensory Sensitivity as a Result of Correlated Multisensory Training" and "Improvement in Unisensory Sensitivity as a Result of Multisensory Associative Learning" would not be parts of the same learning continuum.

and therefore the association needs to be established by repeated exposure to coupled stimuli. The establishment of the association enables the auditory stimulus to enhance the processing of the visual stimulus (and vice versa), thus improving performance in visual detection/discrimination in presence of the coupled stimulus, as described in Section "Improvement in Unisensory Sensitivity as a Result of Multisensory Associative Learning." Once this multisensory association is established, the pairing of the auditory–visual stimuli will not only improve processing at the time of stimulation (as described in Improvement in Unisensory Sensitivity as a Result of Multisensory Associative Learning) but will also lead to plasticity within and between the sensory representations of these associated features, producing the facilitation and enhancement that occurs in the absence of multisensory stimulation, as described in Section "Improvement in Unisensory Sensitivity as a Result of Correlated Multisensory Training." This could be the result of visual and multisensory representations eventually becoming equivalent, where exposure to a unisensory stimulus could invoke the multisensory representation, without the need for multisensory stimulation. Such a phenomenon would result in the performance in the visual-alone and auditory–visual conditions to become equivalent, as was observed in our study (**Figure 1B**).

While we hypothesize that newly learned multisensory associations can lead to facilitation of learning, it may be the case that repeated pairing of arbitrary auditory and visual stimuli may not be sufficient to lead to lasting enhancement of unisensory processing in the absence of the crossmodal signal. It is possible that this multisensory facilitation of unisensory learning is only possible for auditory and visual features that are ecologically associated, such as auditory and visual motion, or lip movements and voice, etc. These ecologically valid associations may be distinct due to hardwired connectivity in the brain, or learning of synaptic structures that are only possible during the critical period, and no longer possible in the mature brain. If so, then regardless of the amount of exposure, arbitrary auditory and visual features will never progress to the stage of enhanced unisensory processing in the absence of the coupled stimulus, and the phenomena discussed in Sections "Improvement in Unisensory Sensitivity as a Result of Correlated Multisensory Training" and "Improvement in Unisensory Sensitivity as a Result of Multisensory Associative Learning" represent two separate learning phenomena as opposed to stages of the same learning continuum. Further research is required to address these questions and to shed light on the neural and computational mechanisms mediating the three types of phenomena outlined in this paper.

We conclude that experience with multisensory stimulus arrays can have a profound impact on processing of unisensory stimuli. This can be through instant recalibrations of sensory maps (see Change in Unisensory Map as a Result of Exposure to Crossmodal Error), the formation of new linkages between auditory and visual features (see Improvement in Unisensory Sensitivity as a Result of Multisensory Associative Learning), or the unisensory representations becoming increasingly indistinct from multisensory representations (see Improvement in Unisensory Sensitivity as a Result of Correlated Multisensory Training). While these are operationally distinct processes, we suggest that there are linkages between the three. For example, enhancement of unisensory representations as well as recalibration of sensory maps both require establishment of their association. While further research will be required to better understand each of these types of learning, and how they relate to each other, it is now clear that the concept of unisensory processing is limited at best, and that prior multisensory exposure can affect perception within a single sensory modality even when the immediate inputs being processed are unisensory.

## REFERENCES

Ahissar, M., and Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends Cogn. Sci. (Regul. Ed.)* 8, 457–464.

Alais, D., and Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.* 14, 257–262.

Ball, K., and Sekuler, R. (1982). A specific and enduring improvement in visual motion discrimination. *Science* 218, 697–698.

Ball, K., and Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision Res.* 27, 953–965.

Birsh, J. R. (1999). *Multisensory Teaching of Basic Language Skills.* Baltimore: Paul H. Brookes Publishing Co.

Botvinich, M., and Cohen, J. (1998). Rubber hands "feel" touch that eyes see. *Nature* 391, 756.

Calvert, G., Spence, C., and Stein, B. E. (2004). *The Handbook of Multisensory Processes.* Cambridge: The MIT Press.

Canon, L. K. (1970). Intermodality inconsistency of input and directed attention as determinants of the nature of adaptation. *J. Exp. Psychol.* 84, 141–147.

Chalk, M., Seitz, A. R., and Series, P. (2010). Rapidly learned stimulus expectations alter perception of motion. *J. Vis.* 10, 1–18.

Coffield, F., Moseley, D., Hall, E., and Ecclestone, K. (2004). *Learning Styles and Pedagogy in Post-16 Learning: A Systematic and Critical Review.* London: Learning and Skills Research Centre.

DeBello, W. M., and Knudsen, E. I. (2004). Multiple sites of adaptive plasticity in the owl's auditory localization pathway. *J. Neurosci.* 24, 6853–6861.

Driver, J., and Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on "sensory-specific" brain regions, neural responses, and judgments. *Neuron* 57, 11–23.

Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *J. Vis.* 7, 7 1–14.

Ernst, M. O., and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 429–433.

Fahle, M. (2004). Perceptual learning: a case for early selection. *J. Vis.* 4, 879–890.

Fahle, M., and Poggio, T. (2002). *Perceptual Learning.* Cambridge, MA: MIT Press.

Feldman, D. E., and Knudsen, E. I. (1997). An anatomical basis for visual calibration of the auditory space map in the barn owl's midbrain. *J. Neurosci.* 17, 6820–6837.

Fernald, G. M., and Keller, H. (1921). The effect of kinesthetic factors in development of word recognition in the case of non-readers. *J. Educ. Res.* 4, 355–377.

Frissen, I., Vroomen, J., De Gelder, B., and Bertleson, P. (2003). The aftereffects of ventriloquism: are they sound-frequency specific? *Acta Psychol. (Amst.)* 113, 315–327.

Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 815–836.

Fujisaki, W., Shimojo, S., Kashino, M., and Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nat. Neurosci.* 7, 773–778.

Ghazanfar, A., and Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends Cogn. Sci. (Regul. Ed.)* 10, 278–285.

Ghose, G. M. (2004). Learning in mammalian sensory cortex. *Curr. Opin. Neurobiol.* 14, 513–518.

Gilbert, C. D., Sigman, M., and Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron* 31, 681–697.

Gingras, G., Rowland, B. A., and Stein, B. E. (2009). The differing impact of multisensory and unisensory integration on behavior. *J. Neurosci.* 29, 4897–4902.

Guo, J., and Guo, A. (2005). Crossmodal interactions between olfactory and visual learning in Drosophila. *Science* 309, 307–310.

Hofma, P. M., Van Riswick, J. G. A., and Van Opstal, A. J. (1998). Relearning sound localization with new ears. *Nat. Neurosci.* 1, 417–421.

Kim, R. S., Seitz, A. R., and Shams, L. (2008). Benefits of stimulus congruency for multisensory facilitation of visual learning. *PLoS ONE* 3, e1532. doi:10.1371/journal.pone.0001532

King, A. (2009). Visual influences on auditory spatial learning. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 331–339.

Lehmann, S., and Murray, M. M. (2005). The role of multisensory memories in unisensory object discrimination. *Brain Res. Cogn. Brain Res.* 24, 326–334.

Lewald, J. (2002). Rapid adaptation to auditory-visual spatial disparity. *Learn. Mem.* 9, 268–278.

Liu, Z. (1999). Perceptual learning in motion discrimination that generalizes across motion directions. *Proc. Natl. Acad. Sci. U.S.A.* 96, 14085–14087.

Miyazaki, M., Yamamoto, S., Uchida, S., and Kitazawa, S. (2006). Bayesian calibration of simultaneity in tactile temporal order judgment. *Nat. Neurosci.* 9, 875–877.

Montessori, M. (1912). *The Montessori Method.* New York: Frederick Stokes.

Montessori, M. (1967). *The Absorbant Mind.* New York: Holt, Rinehart and Winston.

Murray, M. M., Foxe, J. J., and Wylie, G. R. (2005). The brain uses single-trial multisensory memories to discriminate without awareness. *Neuroimage* 27, 473–478.

Murray, M. M., Michel, C. M., Grave De Peralta, R., Ortigue, S., Brunet, D., Gonzalez Andino, S., and Schnider, A. (2004). Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. *Neuroimage* 21, 125–135.

Navarra, J., Hartcher-O'Brien, J., Piazza, E., and Spence, C. (2009). Adaptation to audiovisual asynchrony modulates the speeded detection of sound. *Proc. Natl. Acad. Sci. U.S.A.* 106, 9169–9173.

Newell, F. N., Ernst, M. O., Tjan, B. S., and Bülthoff, H. H. (2001). Viewpoint dependence in visual and haptic objects recognition. *Psychol. Sci.* 12, 37–42.

Orton, S. T. (1928). Specific reading disability – strephosymbolia. *JAMA* 90, 1095–1099.

Pilly, P. K., Grossberg, S., and Seitz, A. R. (2010). Low-level sensory plasticity during task-irrelevant perceptual learning: evidence from conventional and double training procedures. *Vision Res.* 50, 424–432.

Radeau, M., and Bertelson, P. (1974). The after-effects of ventriloquism. *Q. J. Exp. Psychol.* 26, 63–71.

Rao, R. P., and Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, 79–87.

Recanzone, G. (1998). Rapidly induced auditory plasticity: the ventriloquism aftereffect. *Proc. Natl. Acad. Sci. U.S.A.* 95, 869–875.

Recanzone, G. H. (2009). Interactions of auditory and visual stimuli in space and time. *Hear. Res.* 258, 89–99.

Schroeder, C. E., and Foxe, J. (2005). Multisensory contributions to low-level, "unisensory" processing. *Curr. Opin. Neurobiol.* 15, 454–458.

Seitz, A. R., and Dinse, H. R. (2007). A common framework for perceptual learning. *Curr. Opin. Neurobiol.* 17, 148–153.

Seitz, A. R., Kim, R., and Shams, L. (2006a). Sound facilitates visual learning. *Curr. Biol.* 16, 1422–1427.

Seitz, A. R., Nanez, J. E., Holloway, S. R., Tsushima, Y., and Watanabe, T. (2006b). Two cases requiring external reinforcement in perceptual learning. *J. Vis.* 6, 966–973.

Shams, L., and Seitz, A. (2008). Benefits of multisensory learning. *Trends Cogn. Sci. (Regul. Ed.)* 12, 411–417.

Shimojo, S., and Shams, L. (2001). Sensory modalities are not separate modalities: plasticity and interactions. *Curr. Opin. Neurobiol.* 11, 505–509.

Strauss, A., and Lehtinen, L. E. (1947). *Psychopathology and Education of the Brain-Injured Child.* New York: Grune & Stratton.

Sumby, W. H., and Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *J. Acoust. Soc. Am.* 26, 212–215.

Thompson, V., and Paivio, A. (1994). Memory for pictures and sounds: independence of auditory and visual codes. *Can. J. Exp. Psychol.* 48, 380–396.

Treichler, T. G. (1967). Are you missing the boat in training aid? *Film AV Commun.* 1, 14–16.

Von Kriegstein, K., and Giraud, A. L. (2006). Implicit multisensory associations influence voice recognition. *PLoS Biol.* 4, e326. doi:10.1371/journal.pbio.0040326

Wozny, D. R., Seitz, A. R., and Shams, L. (2008). Learning associations between simple visual and auditory features. *J. Vis.* 8, 171.

Wozny, D. R., and Shams, L. (2011). Recalibration of auditory space following milliseconds of cross-modal discrepancy. *J. Neurosci.* 31, 4607–4612.

Zwiers, M. P., Van Opstal, A. J., and Paige, G. D. (2003). Plasticity in human sound localization induced by compressed spatial vision. *Nat. Neurosci.* 6, 175–181.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 09 July 2011; accepted: 23 September 2011; published online: 18 October 2011.*

*Citation: Shams L, Wozny DR, Kim R and Seitz A (2011) Influences of multisensory experience on subsequent unisensory processing. Front. Psychology* ***2****:264. doi: 10.3389/fpsyg.2011.00264*

*This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.*

frontiers in
**PSYCHOLOGY**

**ORIGINAL RESEARCH ARTICLE**
published: 13 April 2011
doi: 10.3389/fpsyg.2011.00055

# Cross-sensory facilitation reveals neural interactions between visual and tactile motion in humans

***Monica Gori*[1]*, *Giacomo Mazzilli*[2], *Giulio Sandini*[1] and *David Burr*[3]**

[1] *Istituto Italiano di Tecnologia, Robotics, Brain and Cognitive Sciences, Genova, Italy*
[2] *School of Psychology, University of Birmingham, Birmingham, UK*
[3] *Dipartimento di Psicologia, Università Degli Studi di Firenze, Florence, Italy*

***Edited by:***
*Nadia Bolognini, University of Milano-Bicocca, Italy*
*Angelo Maravita, University of Milano-Bicocca, Italy*

***Reviewed by:***
*Alberto Gallace, University of Milano-Bicocca, Italy*
*Claudio Brozzoli, Karolinska Institutet, Sweden*

****Correspondence:***
*Monica Gori, Department of Robotics, Brain and Cognitive Science, Italian Institute of Technology, Genoa, Italy.*
*e-mail: monica.gori@iit.it*

Many recent studies show that the human brain integrates information across the different senses and that stimuli of one sensory modality can enhance the perception of other modalities. Here we study the processes that mediate cross-modal facilitation and summation between visual and tactile motion. We find that while summation produced a generic, non-specific improvement of thresholds, probably reflecting higher-order interaction of decision signals, facilitation reveals a strong, direction-specific interaction, which we believe reflects sensory interactions. We measured visual and tactile velocity discrimination thresholds over a wide range of base velocities and conditions. Thresholds for both visual and tactile stimuli showed the characteristic "dipper function," with the minimum thresholds occurring at a given "pedestal speed." When visual and tactile coherent stimuli were combined (summation condition) the thresholds for these multisensory stimuli also showed a "dipper function" with the minimum thresholds occurring in a similar range to that for unisensory signals. However, the improvement of multisensory thresholds was weak and not directionally specific, well predicted by the maximum-likelihood estimation model (agreeing with previous research). A different technique (facilitation) did, however, reveal direction-specific enhancement. Adding a non-informative "pedestal" motion stimulus in one sensory modality (vision or touch) selectively lowered thresholds in the other, by the same amount as pedestals in the same modality. Facilitation did not occur for neutral stimuli like sounds (that would also have reduced temporal uncertainty), nor for motion in opposite direction, even in blocked trials where the subjects knew that the motion was in the opposite direction showing that the facilitation was not under subject control. Cross-sensory facilitation is strong evidence for functionally relevant cross-sensory integration at early levels of sensory processing.

**Keywords: visual, tactile, motion, perception, neural, cross-sensory, interactions, facilitation**

## INTRODUCTION

### INTEGRATION AND INTERACTIONS BETWEEN TOUCH AND VISION

The different sensory modalities provide redundant information about the environment. Much evidence over the last decade has shown that our nervous system integrates signals from different modalities to maximize the information available for perception and action (e.g., Ernst and Bulthoff, 2004). However, it is not entirely clear whether these processes occur at low sensorial levels or higher decisional levels.

For example, much psychophysical evidence demonstrates integration of visual and auditory motion stimuli, but the integration would not be functionally useful in discrimination direction of motion, of a common object, as it occurs equally for both the same and opposite directions of motion (Meyer and Wuerger, 2001; Wuerger et al., 2003; Alais and Burr, 2004a). This suggests that it is not the sensory, directional signals that are being integrated, but "decision signals," unsigned signals that motion has occurred.

On the other hand, other studies point clearly to neural interactions between the senses, using electrophysiological and imaging techniques (e.g., Bolognini and Maravita, 2007; Nakashita et al., 2008; Wang et al., 2008; Kayser et al., 2009; Beauchamp et al., 2010). It is particularly interesting that many areas thought to be primary sensory areas have been showed to respond to several modalities: for example, MT, an early visual area, responds both to tactile and to auditory motion (e.g., Hagen et al., 2002; Blake et al., 2004; Beauchamp et al., 2007; Saenz et al., 2008; Bedny et al., 2010). Similarly, area SI and SII, primary somatosensory areas respond well to visual stimuli (Keysers et al., 2004; Blakemore et al., 2005; Schaefer et al., 2005) and neurophysiological evidence support the presence of multimodal interactions at the level of the single neuron (e.g., Stein et al., 2001; Rowland and Stein, 2007).

That many sensory modalities respond to the motion of objects renders it an ideal stimulus to investigate intermodal interactions and multisensory integration. Much is known about visual motion, but the workings of tactile and auditory motion remain more obscure. Nevertheless, recent evidence suggests that they share much in common (e.g., Pei et al., 2011). For example, tactile motion is subject to similar illusions observed with visual motion, including a motion flow after-effect (Watanabe et al., 2007), the "aperture problem," the Ouchi illusion (Bicchi et al., 2008) and the "Ternus effect" (Harrar and Harris, 2007).

Several studies show cross-modal interactions for motion perception: visual motion can influence the apparent speed of tactile motion (Bensmaia et al., 2006; Craig, 2006), and also influence the speed and direction of audio motion (Mays and Schirillo, 2005; Lopez-Moliner and Soto-Faraco, 2007). Similarly, auditory motion stimuli affect the direction of tactile motion (Soto-Faraco et al., 2004), mutual enhancement between vision and auditory motion has been observed (Strybel and Vatakis, 2004), and reaction times are fast to visuo-tactile motion then to either alone (Ushioda and Wada, 2007). Integration has also been shown for visuo-tactile apparent motion across fingers, but the integration seemed to occur at a high rather than low level (Harrar et al., 2008). More interestingly, recent evidence reports motion bidirectional aftereffects between vision and touch (Konkle et al., 2009) and also between vision and audition (Kitagawa and Ichihara, 2002), both demonstrating intermodal adaptation.

Many studies, using electrophysiological and imaging techniques, including but not limited to those outlined above, suggest that visual, tactile, and auditory motion perception share common neural mechanisms. However, to date no psychophysical studies have demonstrated a clear, functionally useful interaction between motion signals of different modalities in discriminating the direction of motion of a common object. Indeed, also if some psychophysical works reported facilitation between modalities (Wuerger et al., 2003; Alais and Burr, 2004a) the effect of most of them were small (with exception of Arrighi et al., 2009 for visual–auditory integration in biological motion) and, importantly, unspecific for direction. Perhaps the reason for this is that the studies have not been optimized to reveal neural connectivity. In this study we use two psychophysical techniques, summation and facilitation, and show that they are affected differently by visuo-tactile interactions.

## SUMMATION AND FACILITATION

Two common techniques are used to study vision: summation and facilitation. Although they share much in common (see, for example, Pelli, 1987), the techniques are distinct. Summation (first introduced by Rentschler and Fiorentini, 1974) occurs when two different signals (of the same or different modality), individually below threshold combine to reach threshold. In a two-alternative forced choice paradigm the two signals are displayed together in one interval, which has to be discriminated from the blank interval: both are informative about which interval contains the signals. In the facilitation paradigm, on the other hand, a non-informative "pedestal" is displayed to both intervals; and although it provides no direct evidence of which interval contains the signal, it in fact can increase sensitivity to the test when it is of appropriate strength.

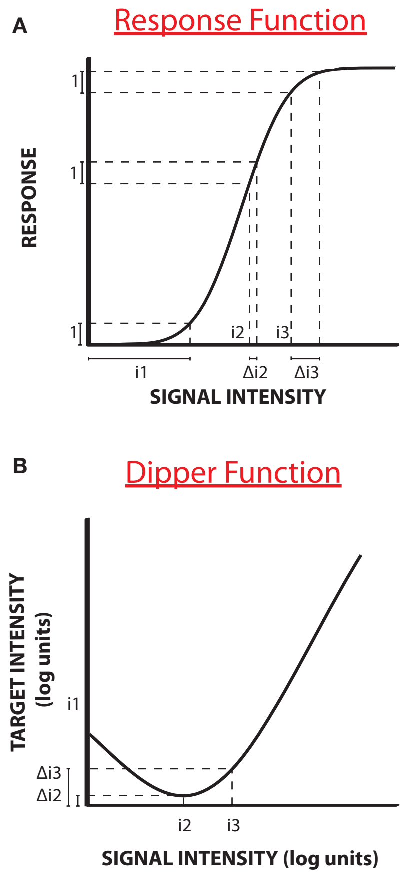

**FIGURE 1 | Response function and dipper function. (A)** A stylized "dipper function": adding a small quantity of signal improves thresholds, but higher levels of base intensity increases them. **(B)** The dip is usually explained by a purported non-linearity in the function that transduces signals into neural responses. The function has an initial threshold-like, accelerating non-linearity and a later compressive non-linearity. If we assume that to perceive two stimuli to be different requires a constant change in responsiveness (illustrated by the dotted lines), then more contrast is required to achieve this change at the lower and higher ends of the curve than in the middle, where the curve is steepest: the lower contrasts at the steep part of the curve cause the dip.

Facilitation is different in that one signal is not informative for detection, but facilitates the detection of the other. It was first demonstrated for luminance and contrast discrimination (Legge and Foley, 1980). **Figure 1A** shows a typical stylized example of facilitation, with a curve plotting threshold of some quantity (such as contrast increment) against the base intensity (say contrast). Nevertheless, the results describe the characteristic "dipper function": adding a small quantity of signal improves thresholds (termed "facilitation"), but higher levels of base intensity increase thresholds (see Solomon, 2009 for a recent review). As the base contrast is not in itself informative about the test, the change in thresholds must reflect a non-linearity in the system. Typically the non-linearity is thought to occur in the function that transduces signals into neural responses (schematically illustrated in **Figure 1B**). The function is thought to show two strong non-linearities, an initial threshold-like accelerating non-linearity and a later compressive non-linearity (e.g., Legge and Foley, 1980). If we assume that to perceive two stimuli as different requires a constant change in responsiveness (illustrated by the dotted lines), then more contrast is required to achieve this change at the lower and higher ends of the curve than in the middle, where the curve is steepest: the lower contrasts at the steep part of the curve cause the dip. Many discrimination functions exhibit a "dipper function" both within the visual modality for example in the discrimination of contrast (Nachmias and Kocher, 1970; Pelli, 1985), blur (Watt and Morgan, 1983; Burr and Morgan, 1997) and motion (Simpson and Finsten, 1995) and between modalities for different functions (Arabzadeh et al., 2008; Burr et al., 2009).

In this study we use both techniques, summation and facilitation, to study interactions between visual and tactile motion. We find that while summation produces a generic, non-specific improvement of thresholds, probably reflecting higher-order interaction of decision signals, facilitation reveals a strong, direction-specific interaction, which we believe reflects sensory interactions. These data have been presented at the visual science conference in Naples 2008 and published in an abstract form (Gori et al., 2008).

## MATERIALS AND METHODS

With the summation technique we studied visual, tactile, and bimodal visuo-tactile motion perception by measuring minimal speed increment motion thresholds over a wide range of base speeds (from 0 to 25 cm/s). The stimuli used were physical wheels etched with a sinewave profile of 10 and 3.3 cycles/cm (**Figures 2A,B**). They were spatially aligned to give the appearance of a common object and simultaneously driven at specific speeds by two independent computer controlled motors (**Figures 2C,D**). Subjects, seated at 57 cm, observed the front wheel and touched the second wheel (concealed from view **Figures 2E,F**) with their index finger. Speed detection and discrimination thresholds were measured with a 2IFC procedure. Subjects chose which of two presentations seemed faster: one presentation (randomly first or second) was the standard, fixed at a specific velocity for each session, while the other stimulus was slightly faster, with speed chosen by the QUEST algorithm (Watson and Pelli, 1983) that homed in near threshold (for each condition 150 trials were collected).

Data were fitted with a cumulative Gaussian function and SEs in the thresholds were computed with bootstrap simulation (Efron and Tibshirani, 1993). In the "vision only" task the visual stimulus

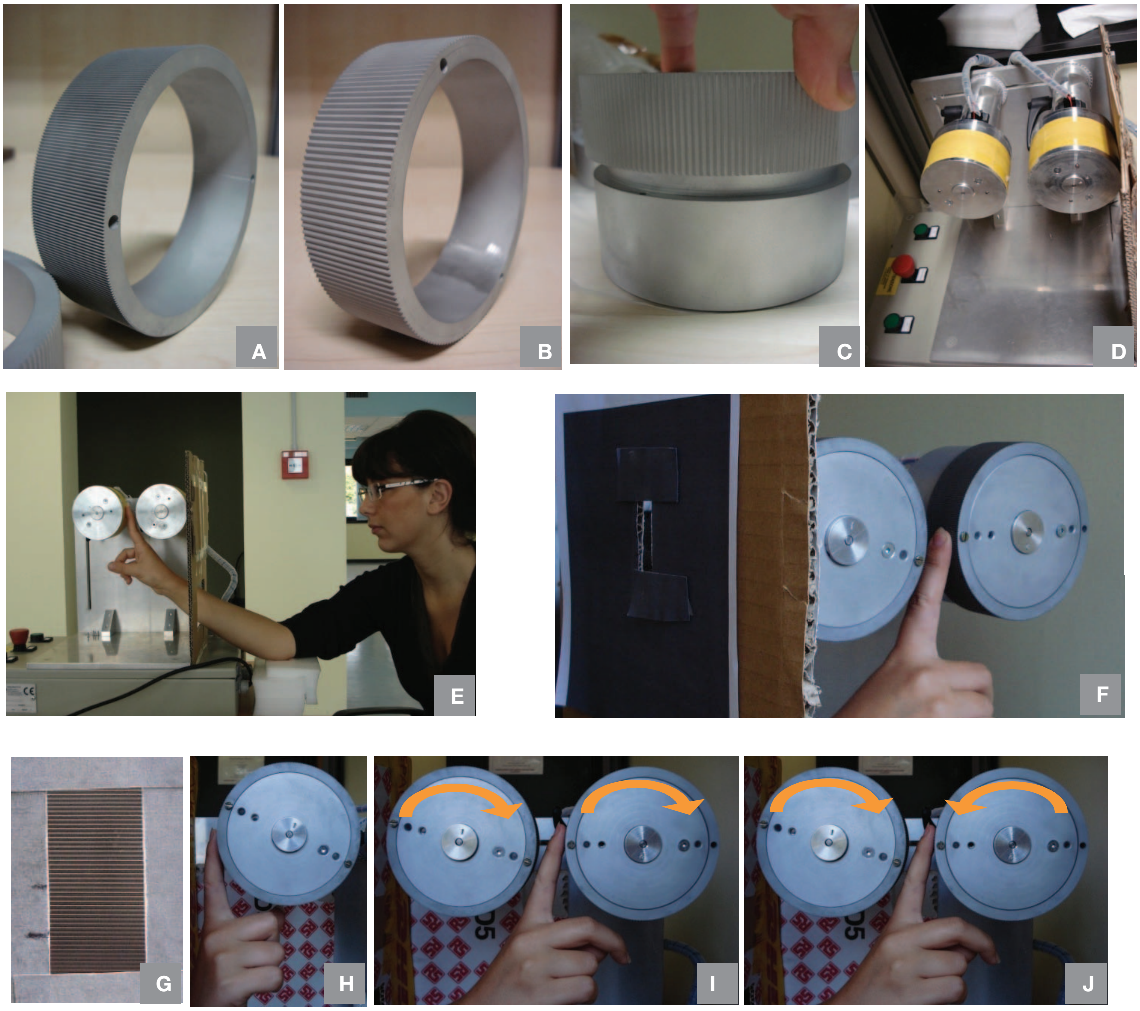

**FIGURE 2 | Stimuli: physical wheels etched with a sinewave profile of different spatial frequencies: (A) 10 cycles/cm and a phase of 1 mm, (B) 3.3 cycles/cm, and a phase of 3 mm. (C)** Support where the wheels were inserted. **(D)** Setup with two arms driven at variable speeds by two independent computer controlled motors. **(E)** Example subject during the experiment. **(F)** Closeup of the visual-haptic stimulation unit. **(G)** Image of the "vision only" stimulus. **(H)** Image of the "tactile only" stimulus. **(I)** Image of the "bimodal task same direction of motion." **(J)** Image of the "bimodal task opposite direction of motion."

was viewed through a small window (1 × 3 cm **Figure 2G**). For the entire presentation the subject had to observe a fixation point. In the "tactile only" task the subject touched the tactile stimulus with the fingertip of his index finger (1 × 2 cm **Figure 2H**). During the "bimodal task" (**Figure 2I**) subjects were instructed to touch and observe simultaneously the two wheels moving in the same direction. During the "bimodal, opposite direction" task (**Figure 2J**) the two wheels were moved in opposite directions. To control that the final velocity and acceleration profile of the wheels was equal to the one required by the experimenter, we recorded the velocity profile of the wheels in motion by using a motion tracking system (Optotrack Certus system). For all the considered stimuli the measured speed profile was consistent with the requirements. The maximum time required to the wheel to reach the maximal final velocity with the maximal acceleration used in our experiment was 0.0057 s.

With the facilitation technique only the 10 cycles/cm stimuli were used (**Figure 2A**). In the cross-modal pedestal condition, the base speed (*pedestal*) was presented in one modality, and the increment to be detected in the other. The facilitation effect was measured for different speeds of the pedestal stimulus from ranging from 0 to 10 cm/s. During the visual detection plus the tactile pedestal signal, the subject had to perform a visual detection task and at the same time, for both the intervals, he was stimulated by a tactile signal driven at a specific pedestal speed equal for both the presentations. During the tactile detection plus visual pedestal signal, the subject had to perform a tactile detection task and had to observe at the same time, for both intervals, a visual stimulus driven at a specific pedestal speed for both the intervals. For example, in both intervals the tactile motion could be, say, 1 cm/s (hence non-informative), and only in the *test* interval was there the visual motion to be detected, and the threshold of that motion was detected. We then performed some control experiments by measuring the facilitation effect for pedestal signals of different origins. In the first control experiment, we substituted the cross-modal motion-pedestal with a sound of matched duration (defining precisely the temporal interval of motion). The subject had to perform the same unimodal visual and tactile detection tasks but during each stimulation he was presented with a synchronized acoustic signal. In the second control experiment, we measured facilitation with cross-modal pedestals moving in the opposite direction to the tests and the subject was informed from the experimenter about the future direction of the two motions. In the third control experiment we delayed the beginning of the cross-modal pedestal stimulus of 100 ms with respect to the test stimulus. Two subjects took part to all the experiments of this study. The most important data, however, were collected on five subjects (indicated in figure captions).

## RESULTS

### WITHIN MODALITY FACILITATION

**Figure 3** reports discrimination and detection thresholds in two subjects for visual (on the left) and tactile (on the right) stimuli for two spatial frequencies (3.3 and 10 cycles/cm). Both visual and tactile motion produced the characteristic "dipper function," where the thresholds initially decreased with base speed to a minimum at base speeds around 0.1 cm/s, then rose, roughly in proportion to base speed (Weber law behavior). The visual and tactile motion curves are very similar in form, both in absolute sensitivity and position of the dip, suggesting that similar mechanisms may operate for these two modalities.

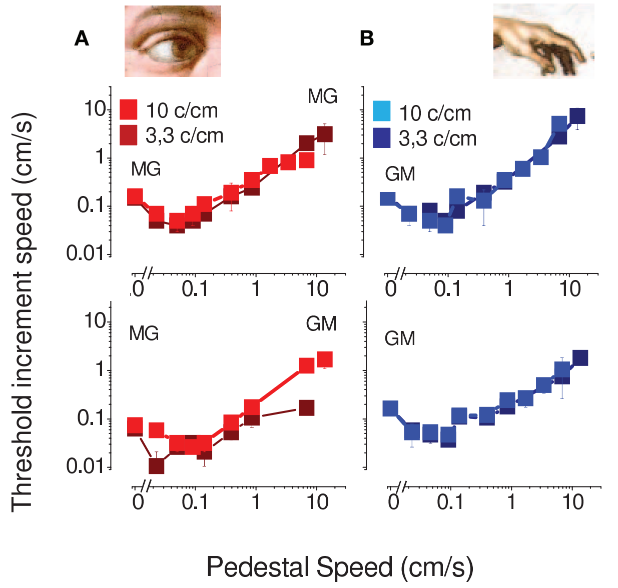

**FIGURE 3 | Speed thresholds in the within modality facilitation condition. (A)** Vision only condition. The spatial frequency of 3.3 cycles/cm is reported light red and the spatial frequency of 10 cycles/cm dark red. **(B)** Tactile only condition. The spatial frequency of 3.3 cycles/cm is reported light blue and the spatial frequency of 10 cycles/cm dark blue. This measure was replicated in two subjects.

### BETWEEN MODALITY SUMMATION

**Figure 4** plots the data in another way, separately for the two spatial frequencies (3.3 and 10 cycles/cm). All curves, for visual, tactile, and visuo-tactile motion (respectively red, blue, and green symbols) are similar in form, with the "dip" (maximum facilitation) always around 0.1 cm/s. Interestingly, for multisensory motion, the functions (including the size and position of the "dip") were virtually identical when the visual–tactile motion were driven in the same or opposite direction (respectively green and violet symbols). We modeled the predicted improvements for multisensory stimuli using the standard maximum-likelihood model (Ernst and Banks, 2002; Alais and Burr, 2004b). Optimal integration for the visual–tactile stimulus ($\sigma_{VT}$) is given by:

$$\sigma_{VT}^2 = \frac{\sigma_V^2\sigma_T^2}{\sigma_V^2+\sigma_T^2} \le \min\left(\sigma_V^2,\sigma_T^2\right) \qquad (1)$$

Where $\sigma_V$ and $\sigma_T$ are the visual and tactile thresholds. Both the bimodal conditions were well predicted by this model (orange curves in **Figure 4**) in all conditions, suggesting that visuo-tactile flow signals are integrated in a Bayesian optimal fashion.

We measured two points of the curves (for the spatial frequency of 10 cycles/cm) – visual and tactile only thresholds plus points of maximum facilitation– in five subjects (**Figure 4C**). We also measured the thresholds of five subjects by adding a higher

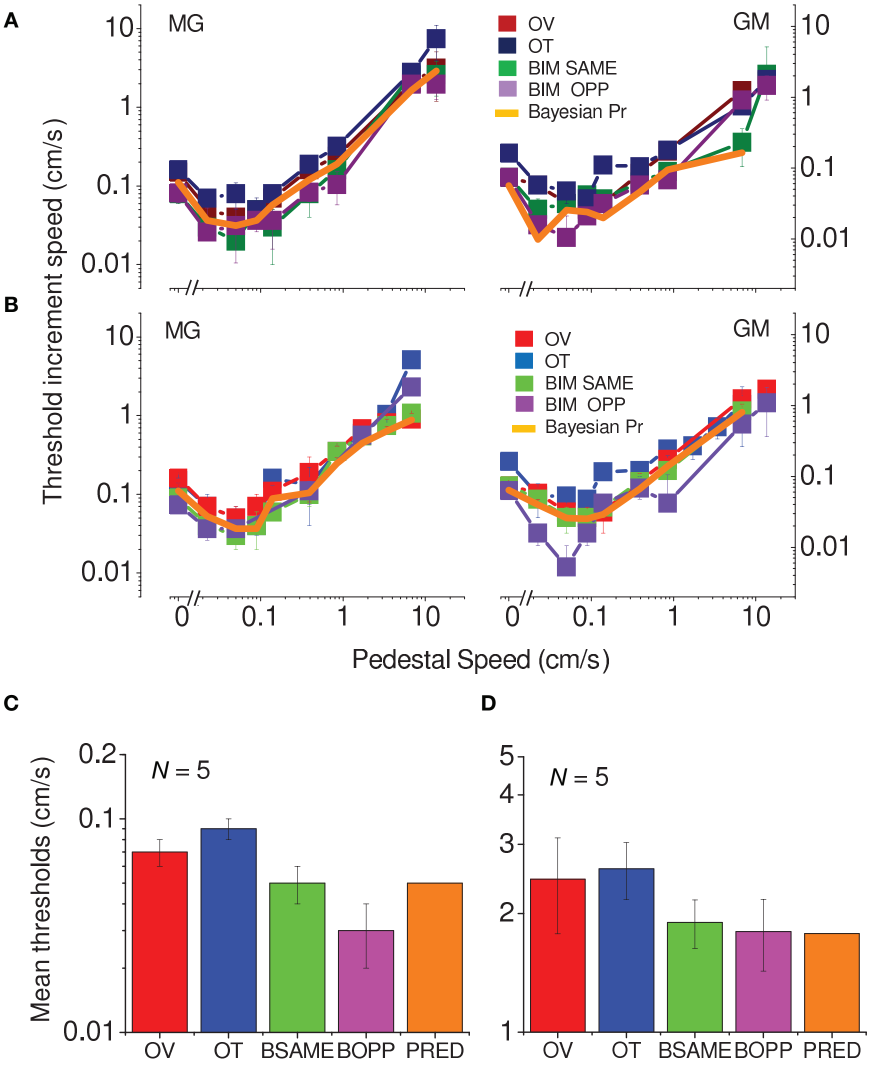

**FIGURE 4 | Speed thresholds for summation between modalities.** **(A)** Spatial frequency of 3.3 cycles/cm. Thresholds for vision only are reported in dark red, for tactile only in dark blue and for bimodal in dark green. Thresholds for bimodal same direction condition light green, for bimodal opposite direction light violet, Bayesian prediction orange (replicated in two subjects). **(B)** Spatial frequency of 10 cycles/cm. Thresholds for vision only are reported in light red, tactile only in light blue, bimodal same direction condition light green, bimodal opposite direction light violet, Bayesian prediction in orange (replicated in two subjects). **(C)** Average thresholds of five subjects for the velocity of 0.1 cm/s (color-coding as above). Spatial frequency of 10 cycles/cm. **(D)** Average thresholds of five subjects for the velocity of 7.5 cm/s (color-coding as above). Spatial frequency of 10 cycles/cm.

velocity (**Figure 4D**). All subjects showed the same effect: a decrease in thresholds for visuo-tactile motion, both in the same and opposite directions; and all were well predicted by the Bayesian maximum-likelihood model (orange bars).

## FACILITATION BETWEEN SENSES

The previous results demonstrate facilitation within the visual and tactile modalities, and also show that the two modalities summate with each other, but in a non-specific manner. Here we ask whether visual motion can facilitate tactile motion, and *vice versa*. Subjects were asked to discriminate which interval contained the visual motion, when in both intervals there was tactile motion, and *vice versa*. The results are shown in **Figure 5** with blue symbols, together with the previously reported results (red symbols). As with within-sensory facilitation, the cross-sensory data show a clear "dip," again at around 0.1 cm/s, clearly demonstrating facilitation across senses. However, the form of the blue curves differed from the within-sensory curves in that there is no rising limb with Weber-like behavior.

To study the facilitation more closely, we measured visual and tactile speed thresholds with and without 0.1 cm/s pedestals of the same or different modality in five subjects. All subjects showed a strong facilitation when a pedestal signal of the same modality was added to the original signal (red symbols in **Figure 6C**) or different modality (blue symbols in **Figure 6C**). Averaged results (normalized to base threshold) are shown in **Figures 6A,B**. For both vision and touch, pedestals of the same (red bars) or different (blue) modality both reduced thresholds considerably, by more than a

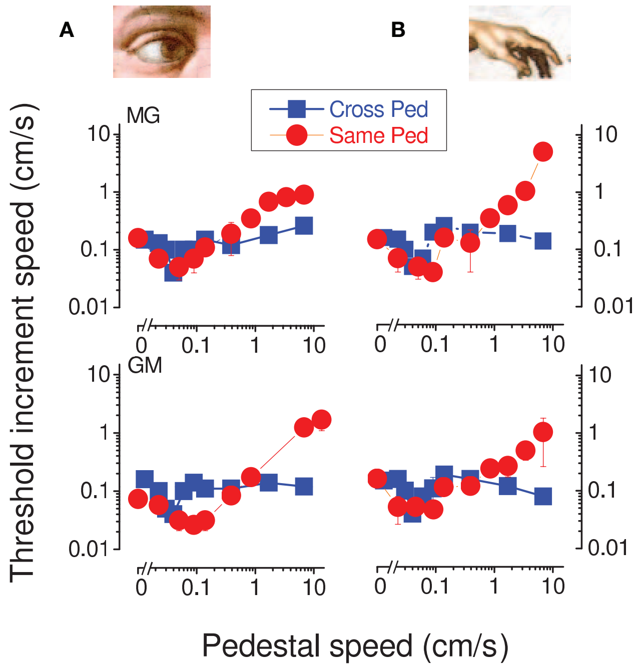

**FIGURE 5 | Incremental speed thresholds for the facilitation between senses condition incremental speed thresholds for visual (A) and tactile (B) motion, as a function of base speed, for two observers.** Red circular symbols show unimodal thresholds, where all signals are confined to the same modality, vision **(A)** or touch **(B)**. The blue symbols show thresholds for pedestals of different modality to the test, tactile **(B)** or visual **(A)**.

factor of two. In both cases the average effect of the pedestal was as strong for the cross-modal as for the intra-modal condition. To examine whether this may be due to reducing *temporal uncertainty*, we substituted the cross-modal motion-pedestal with a sound of matched duration (defining precisely the temporal interval of motion); but the concurrent sounds had no effect on base thresholds (green bars). More importantly, we also measured facilitation with cross-modal pedestals moving in the opposite direction to the tests (and subjects were informed that this was the case), but this had no effect on base thresholds (cyan bars). **Figure 6C** shows the individual data for all conditions, plotting the pedestal against the no-pedestal thresholds. Clearly all red and blue symbols (pedestal conditions of the same or crossed modality) lie under the equality line, showing facilitation, while the other two conditions are at or above it, showing no facilitation.

Uncertainty affects the psychometric function in several ways, most notably in that high uncertainty causes a steepening of the psychometric function (Pelli, 1985; Henning and Wichmann, 2007). **Figure 7** reports the mean normalized steepness of the psychometric function of all measured conditions. As is apparent, all the conditions with a pedestal or accompanying sound have a broader function than the baseline, single modality conditions, with no specific effect for those conditions that lead to an improvement of thresholds, making uncertainty an unlikely explanation for the facilitation effect. In a third control experiment we measured (in three subjects) the facilitation by delaying the beginning of the cross-modal pedestal stimulus of 100 ms with respect to the test stimulus. This short delay was sufficient to disrupt completely the effect of facilitation, suggesting that the test and pedestal had to be simultaneous for facilitation to occur.

## DISCUSSION

We have investigated summation and facilitation of visual and tactile motion, both within and between senses. The results show that the two forms of motion are processed by similar mechanisms, which interact with each other, possibly at an early level of analysis. Over a wide range of speeds, the motion sensitivity curves for both vision and touch are very similar, both showing a dipper-like facilitation at around 0.1 cm/s. When the two modalities were presented together in a *summation* paradigm, sensitivity improved by a factor of about root-two, the amount predicted by the standard maximum-likelihood model. However, the improvement was non-specific for motion direction, as previously reported for audio–visual motion (Meyer and Wuerger, 2001; Wuerger et al., 2003; Alais and Burr, 2004a). Non-specific summation of this sort (combining upward visual with downward tactile motion) would have little functional advantage to perception as opposite motion signals cannot arise from a single object, and probably reflects a mere statistical advantage of having two rather than one signal. This process, traditionally termed "probability summation" (e.g., Watson, 1979; Pelli, 1985; Graham, 1989), basically reflects the increased probability that at least one of two noisy signals will be detected. The predicted magnitude of the effects depends on the slope of the psychometric functions but basically is near what we observe here (in the order of root-two). The most interesting result reported here is the clear, directional-specific cross-sensory facilitation between tactile and visual motion. The effect was larger than that observed for summation, a two- or three-fold increase in sensitivity (compared with the root-two summation effect) and, most importantly, occurred only for motion in the same direction presented at the same time similar to what Arrighi et al. (2009) showed with "tap dancing." The specificity of the interaction suggests that it is functionally important, allowing the system to combine signals from the two senses generated by a common objects to detect motion with a higher sensitivity than with one alone. There are two strong indications that the interactions occur at a low, sensory level rather than at a higher cognitive level. Firstly, when we measured facilitation with motion in the opposite direction, subjects were informed that the tactile and visual motion were in opposite directions. In theory, subjects could have cognitively inverted the motion and taken advantage of it in the same way as motion in the same direction, but this did not happen. Similarly, a delayed pedestal had no effect on thresholds, suggesting that a sensory interaction was necessary.

Discrimination functions for many tasks follow a "dipper function" (see Solomon, 2009 for a recent review), including contrast discrimination (Nachmias and Kocher, 1970; Pelli, 1985), blur (Watt and Morgan, 1983), visual motion (Simpson and Finsten, 1995) and even temporal discrimination (Burr et al., 2009). The usual explanation for the dipper function is that it reflects essential non-linearities in the neural response curve (**Figure 1**; Legge and Foley, 1980; Legge et al., 1987). For tactile motion to affect the visual neural response curve (and *vice versa*) they must interact at a fairly early stage of analysis. Other explanations of the dipper

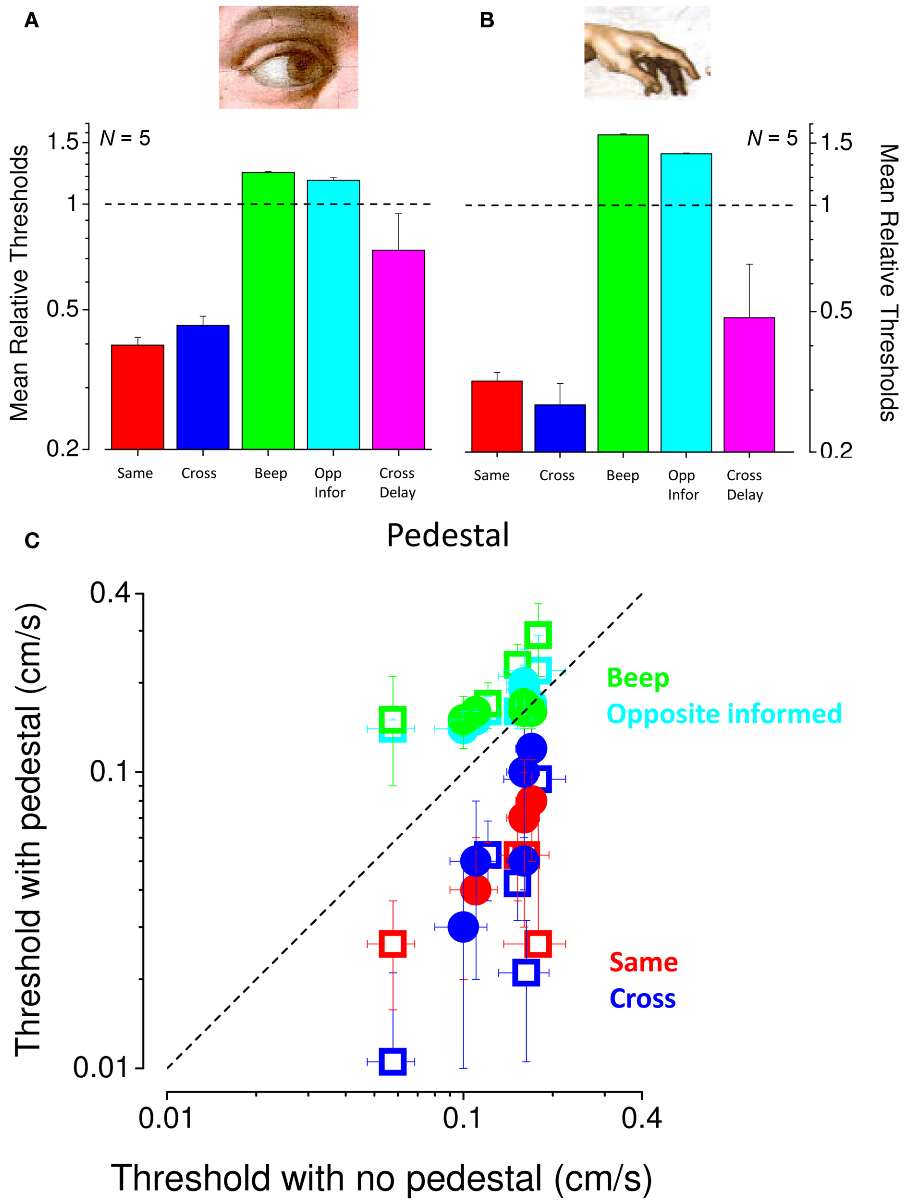

**FIGURE 6 | Conditions producing cross-modal facilitation. (A,B)** Mean normalized thresholds of five observers for visual **(A)** and tactile **(B)** speed increment discrimination. Individual thresholds were divided by thresholds for the no-pedestal condition, then averaged (geometric mean) across subjects (error bars ± 1 SEM). The dashed line at unity indicates no-pedestal effect. Red bars refer to thresholds for pedestals of the same modality, blue pedestals of different modality. The green bars show thresholds when the interval was marked by an auditory tone of 2450 Hz, and the cyan bar thresholds opposite directions of motion (observers were informed of the inversion). Pink bars indicate thresholds for pedestal signal of the other modality delayed of 100 ms (measured in three subjects). **(C)** Individual thresholds for speed increment discrimination, plotted against no-pedestal thresholds. Visual thresholds are shown by closed symbols, tactile by open symbols (color coding as above). Error bars on individual data points were obtained by bootstrap (Efron and Tibshirani, 1993).

function involve spatiotemporal uncertainty (Pelli, 1985), essentially suggesting that the pedestal reduces the time window – and hence the noise within that window – that needs to be monitored. However, the lack of facilitation by sound beeps, or by motion of opposite direction, combined with the fact that slope of the psychometric increased in all dual-modality conditions (irrespective of whether the pedestal caused facilitation) excludes this potential explanation. For facilitation to occur, the motion of the two senses needs to be in the same direction, at the same time, and have a very similar speed: all this points to neural combination that is not under the subject's control.

We have attempted to account for our results within a simple schematic model (**Figure 8**). Taken together, the summation and facilitation results suggest that visual and tactile motion signals interact at least two levels, a relatively low, direction-specific sensory level and a higher-level, aspecific, "probabilistic" interaction that explains

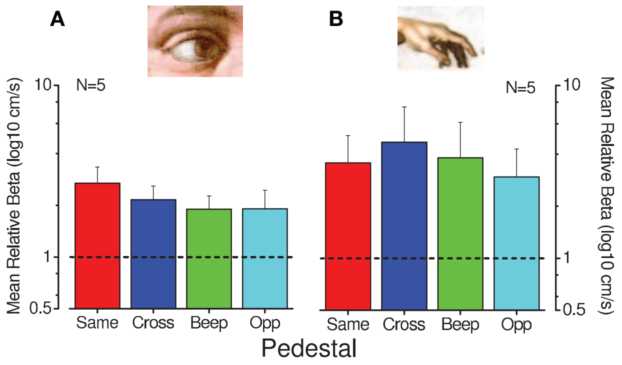

**FIGURE 7 | Steepness of psychometric functions for the conditions that produce cross-modal facilitation.** Mean normalized psychometric function steepness of five observers for visual **(A)** and tactile **(B)** conditions. Individual steepness (SD of psychometric functions) were divided by the steepness for the no-pedestal condition, then averaged across subjects (with error bars representing ± 1 SEM). The dashed line at unity indicates no-pedestal effect: red bars indicate thresholds for pedestals of the same modality, blue pedestals of different modality, green for the auditory tone and cyan for opposite direction of motion.

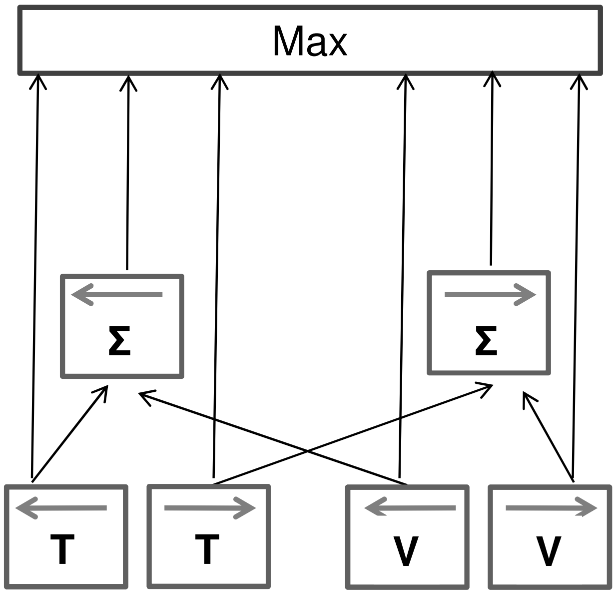

**FIGURE 8 | Descriptive model for the interaction of visual and tactile motion signals.** Visual and tactile motion signals interact at two levels: a low direction-specific sensory level and a higher decisional stage in which a probability interaction occurs. At the first level we assume the signals are vectorially summed (taking direction into account), while at the second level (upper box) we assume a simple maximum operation that chooses the largest signal for each trial. An important aspect of the model is that cross-modal summation is non-compulsory: signals both from the individual senses, and after combination arrive at the decision level. If the combination is less than the individual sensory signals, the maximum operation will choose the unisensory rather than multisensory signals. This schema explains, at least qualitatively, all the observed results.

summation. The facilitation results require that tactile and visual signals are vectorially summed at an early level, taking into account the sign and direction of the motion. However, this summation is clearly not *compulsory*: the system must also have access to signals from the individual senses. There is much evidence for this. For example, the cross-sensory "dipper function" shows no rising phase, where thresholds become worse than in the no-pedestal condition (masking). If the system has access to the individual sensory data, then the "masked" multisensory signal could be ignored when it is weaker than the single-sense signal. Similarly, when test and pedestal move in the opposite directions, there is no cost, as the null combined signal could be ignored in favor of the single-sense signal. To keep the model at its most simplest, we assume that the decision stage (upper box) is a simple "maximum" operation, choosing the largest signal for a particular trial. This "max" operation also explains the small, non-directional summation effects. As the neural signals are "noisy," they vary considerably in velocity from one trial to another: if two signals are present (even if in opposite directions), there is a greater chance that at least one of these will rise above the noisy stationary signal than if only one were present. However, this is a mere statistical advantage, not evidence for neural summation. This model is in agreement with previous results that support the idea of interaction between modalities at a sensorial level (e.g., Ushioda and Wada, 2007; Konkle et al., 2009) but only partially with others that suggest only higher-level multisensory integration (Harrar et al., 2008).

The direction-specific facilitation that we observe suggests that visual and tactile motion share common neural mechanisms. As mentioned earlier, imaging studies have shown that tactile and auditory motion activate several visual cortical areas, including area MT (Hagen et al., 2002; Ricciardi et al., 2004; Saenz et al., 2008), and many studies demonstrate multisensory interactions at the level of the single neurons (see Stein et al., 1993 for a review). This makes MT a highly plausible candidate for the neural substrate underlying the interactions reported here. Interestingly a recent study of Pei et al. (2011) also highlighted that some somatosensory areas and MT have similar functional proprierties, and we can not exclude that both MT and somatosensory areas may be involved in these cross-sensory integration mechanisms. Further fMRI studies under facilitation-like conditions would be usefull in understanding better the neural sustrate for these interactions.

In conclusion, we have shown that the psychophysical technique of *pedestal facilitation* demonstrates clear neural interactions between visual and tactile motion processing. Although neural interactions have been well documented by physiological techniques, psychophysical evidence for these interactions have been elusive to date. This is probably because there are many paths to perception: the senses do combine with each other, and at an early level, but the combination is not obligatory (agreeing with Ernst and Banks, 2002). The system seems to have access to information at various levels, and this parallel access can confound attempts to demonstrate neural intergration. If noisy parallel signals all feed into a simple decision process (such as the absolute maximum schema of **Figure 8**) there will be two important consequences: firstly the parallel access will obscure any negative effects, such as masking or vectorial summation of opposing directions; and secondly it will lead to enhancement of sensitivity for multiple signals, on sheer probabistic grounds, and this enhancement will sometimes obscure small amounts of neural summation. So while our results are certainly consistent with previous results showing aspecific, probabilistic summation, they also show that more appropriate psychophysical techniques can reveal clear, low-level sensory interactions.

## REFERENCES

Alais, D., and Burr, D. (2004a). No direction-specific bimodal facilitation for audiovisual motion detection. *Brain Res. Cogn. Brain Res.* 19, 185–194.

Alais, D., and Burr, D. (2004b). The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.* 14, 257–262.

Arabzadeh, E., Clifford, C. W. G., and Harris, J. A. (2008). Vision merges with touch in a purely tactile discrimination. *Psychol. Sci.* 19, 635–641.

Arrighi, R., Marini, F., and Burr, D. (2009). Meaningful auditory information enhances perception of visual biological motion. *J. Vis.* 9, 1–7.

Beauchamp, M. S., Pasalar, S., and Ro, T. (2010). Neural substrates of reliability-weighted visual-tactile multisensory integration. *Front. Syst. Neurosci.* 4:25. doi: 10.3389/fnsys.2010.00025

Beauchamp, M. S., Yasar, N. E., Fryec, R. E., and Ro, T. (2007). Human MST but not MT responds to tactile stimulation. *J. Neurosci.* 27, 8261–8267.

Bedny, M., Konkle, T., Pelphrey, K., Saxe, R., and Pascual-Leone, A. (2010). Sensitive period for a multimodal response in human visual motion area MT/MST. *Curr. Biol.* 20, 1900–1906.

Bensmaia, S. J., Killebrew, J. H., and Craig, J. C. (2006). Influence of visual motion on tactile motion perception. *J. Neurophysiol.* 96, 1625–1637.

Bicchi, A., Scilingo, E. P., Ricciardi, E., and Pietrini, P. (2008). Tactile flow explains haptic counterparts of common visual illusions. *Brain Res. Bull.* 75, 737–741.

Blake, R., Sobel, K. V., and James, T. W. (2004). Neural synergy between kinetic vision and touch. *Psychol. Sci.* 15, 397–402.

Blakemore, S. J., Bristow, D., Bird, G., Frith, C., and Ward, J. (2005). Somatosensory activations during the observation of touch and a case of vision-touch synaesthesia. *Brain* 128(Pt 7), 1571–1583.

Bolognini, N., and Maravita, A. (2007). Proprioceptive alignment of visual and somatosensory maps in the posterior parietal cortex. *Curr. Biol.* 17, 1890–1895.

Burr, D., Silva, O., Cicchini, G. M., Banks, M. S., and Morrone, M. C. (2009). Temporal mechanisms of multimodal binding. *Proc. Biol. Sci.* 276, 1761–1769.

Burr, D. C., and Morgan, M. J. (1997). Motion deblurring in human vision. *Proc. Biol. Sci.* 264, 431–436.

Craig, J. C. (2006). Visual motion interferes with tactile motion perception. *Perception* 35, 351–367.

Efron, B. T., and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*, New York: Chapman and Hall.

Ernst, M. O., and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 429–433.

Ernst, M. O., and Bulthoff, H. H. (2004). Merging the senses into a robust percept. *Trends Cogn. Sci.* 8, 162–169.

Gori, M., Del Viva, M., Sandini, G., and Burr, D. C. (2008). Young children do not integrate visual and haptic form information. *Curr. Biol.* 18, 694–698.

Graham, N. V. S. (1989). *Visual Pattern Analyzers.* Oxford: Oxford University Press.

Hagen, M. C., Franzen, O., McGlone, F., Essick, G., Dancer, C., and Pardo, J. V. (2002). Tactile motion activates the human middle temporal/V5 (MT/V5) complex. *Eur. J. Neurosci.* 16, 957–964.

Harrar, V., and Harris, L. R. (2007). Multimodal ternus: visual, tactile, and visuo-tactile grouping in apparent motion. *Perception* 36, 1455–1464.

Harrar, V., Winter, R., and Harris, L. R. (2008). Visuotactile apparent motion. *Percept. Psychophys.* 70, 807–817.

Henning, G. B. W., and Wichmann, F. A. (2007). Some observations on the pedestal effect. *J. Vis.* 7, 1–15.

Kayser, C., Logothetis, N. K., and Panzeri, S. (2009). Visual enhancement of the information representation in auditory cortex. *Curr. Biol.* 20, 19–24.

Keysers, C., Wicker, B., Gazzola, V., Anton, J. L., Fogassi, L., and Gallese, V. (2004). A touching sight: SII/PV activation during the observation and experience of touch. *Neuron* 42, 335–346.

Kitagawa, N., and Ichihara, S. (2002). Hearing visual motion in depth. *Nature* 416, 172–174.

Konkle, T., Wang, Q., Hayward, V., and Moore, C. I. (2009). Motion aftereffects transfer between touch and vision. *Curr. Biol.* 19, 745–750.

Legge, G. E., and Foley, J. M. (1980). Contrast masking in human vision. *J. Opt. Soc. Am.* 70, 1458–1471.

Legge, G. E., Kersten, D., and Burges, A. E. (1987). Contrast discrimination in noise. *J. Opt. Soc. Am. A* 4, 391–404.

Lopez-Moliner, J., and Soto-Faraco, S. (2007). Vision affects how fast we hear sounds move. *J. Vis.* 7, 1–7.

Mays, A., and Schirillo, J. (2005). Lights can reverse illusory directional hearing. *Neurosci. Lett.* 384, 336–338.

Meyer, G. F., and Wuerger, S. M. (2001). Cross-modal integration of auditory and visual motion signals. *Neuroreport* 12, 2557–2560.

Nachmias, J., and Kocher, E. C. (1970). Visual detection and discrimination of luminance increments. *J. Opt. Soc. Am.* 60, 382–389.

Nakashita, S., Saito, D. N., Kochiyama, T., Honda, M., Tanabe, H. C., and Sadato, N. (2008). Tactile-visual integration in the posterior parietal cortex: a functional magnetic resonance imaging study. *Brain Res. Bull.* 75, 513–525.

Pei, Y. C., Hsiao, S. S., Craig, J. C., and Bensmaia, S. J. (2011). Neural mechanisms of tactile motion integration in somatosensory cortex. *Neuron* 69, 536–547.

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *J. Opt. Soc. Am. A* 2, 1508–1532.

Pelli, D. G. (1987). On the relation between summation and facilitation. *Vision Res.* 27, 119–123.

Rentschler, I., and Fiorentini, A. (1974). Meridional anisotrophy of psychophysical spatial interactions. *Vision Res.* 14, 1467–1475.

Ricciardi, E., Vanello, N., Dente, D., Sgambellini, N., Scillingo, E. P., Gentili, C., Sani, L., Positano, V., Santarelli, F. M., Guazzelli, M., Haxby, J. V., Landini, L., Bicchi, A., and Pietrin, P. (2004). "Perception of visual and tactile flow activitates common cortical areas in the human brain," in *Proceedings of the EuroHaptics 2004* (Germany: Technische Universitat Munchen), 290–292.

Rowland, B. A., and Stein, B. E. (2007). Multisensory integration produces an initial response enhancement. *Front. Integr. Neurosci.* 1:4. doi: 10.3389/neuro.07/004.2007

Saenz, M., Lewis, L. B., Huth, A. G., Fine, I., and Koch, C. (2008). Visual motion area MT+/V5 responds to auditory motion in human sight-recovery subjects. *J. Neurosci.* 28, 5141–5148.

Schaefer, M., Heinze, H. J., and Rotte, M. (2005). Seeing the hand being touched modulates the primary somatosensory cortex. *Neuroreport* 16, 1101–1105.

Simpson, W. A., and Finsten, B. A. (1995). Pedestal effect in visual motion discrimination. *J. Opt. Soc. Am. A. Opt. Image Sci. Vis.* 12, 2555–2563.

Solomon, J. A. (2009). The history of dipper functions. *Atten. Percept. Psychophys.* 71, 435–443.

Soto-Faraco, S., Spence, C., and Kingstone, A. (2004). Congruency effects between auditory and tactile motion: extending the phenomenon of cross-modal dynamic capture. *Cogn. Affect. Behav. Neurosci.* 4, 208–217.

Stein, B. E., Jiang, W., Wallace, M. T., and Stanford, T. R. (2001). Nonvisual influences on visual-information processing in the superior colliculus. *Prog. Brain Res.* 134, 143–156.

Stein, B. E., Meredith, M. A., and Wallace, M. T. (1993). The visually responsive neuron and beyond: multisensory integration in cat and monkey. *Prog. Brain Res.* 95, 79–90.

Strybel, T. Z., and Vatakis, A. (2004). A comparison of auditory and visual apparent motion presented individually and with crossmodal moving distractors. *Perception* 33, 1033–1048.

Ushioda, H., and Wada, Y. (2007). "Multisensory integration between visual and tactile motion information: evidence from redundant-signals effects on reaction time," in *23rd Annual Meeting of the ISP Proceedings of the 23rd Annual Meeting of the ISP*, Tokyo.

Wang, Y., Celebrini, S., Trotter, Y., and Barone, P. (2008). Visuo-auditory interactions in the primary visual cortex of the behaving monkey: electrophysiological evidence. *BMC Neurosci.* 9, 79. doi: 10.1186/1471-2202-9-79

Watanabe, J., Hayashi, S., Kajimoto, H., Tachi, S., and Nishida, S. (2007). Tactile motion aftereffects produced by appropriate presentation for mechanoreceptors. *Exp. Brain Res.* 180, 577–582.

Watson, A. B. (1979). Probability summation over time. *Vision Res.* 19, 515–522.

Watson, A. B., and Pelli, D. G. (1983). QUEST: a Bayesian adaptive psychometric method. *Percept. Psychophys.* 33, 113–120.

Watt, R. J., and Morgan, M. J. (1983). The recognition and representation of edge blur: evidence for spatial primitives in human vision. *Vision Res.* 23, 1465–1477.

Wuerger, S. M., Hofbauer, M., and Meyer, G. F. (2003). The integration of auditory and visual motion signals at threshold. *Percept. Psychophys.* 65, 1188–1196.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 20 December 2010; accepted: 23 March 2011; published online: 13 April 2011.*

*Citation: Gori M, Mazzilli G, Sandini G and Burr D (2011) Cross-sensory facilitation reveals neural interactions between visual and tactile motion in humans. Front. Psychology* ***2****:55. doi: 10.3389/fpsyg.2011.00055*

*This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.*

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 22 November 2011
doi: 10.3389/fpsyg.2011.00341

# Visuo-haptic interactions in unilateral spatial neglect: the cross modal Judd illusion

***Flavia Mancini*[1,2]*, *Emanuela Bricolo*[1], *Flavia C. Mattioli*[3] and *Giuseppe Vallar*[1,2]***

[1] *Department of Psychology, University of Milano-Bicocca, Milano, Italy*
[2] *Neuropsychological Laboratory, IRCCS Italian Auxological Institute, Milano, Italy*
[3] *Neuropsychological Unit, Spedali Civili di Brescia, Brescia, Italy*

***Edited by:***
*Nadia Bolognini, University Milano-Bicocca, Italy*

***Reviewed by:***
*Tony Ro, City College of New York, USA*
*David Pitcher, Massachusetts Institute of Technology, USA*

****Correspondence:***
*Flavia Mancini and Giuseppe Vallar, Department of Psychology, University of Milano-Bicocca, Piazza dell'Ateneo Nuovo 1, 20126 Milano, Italy.*
*e-mail: f.mancini2@campus.unimib.it; giuseppe.vallar@unimib.it*

Unilateral spatial neglect (USN) has been mainly investigated in the visual modality; only few studies compared spatial neglect across different sensory modalities, and explored their multisensory interactions, with controversial results. We investigated the integration between vision and haptics, through a bisection task of a cross modal illusion, the Judd variant of the Müller-Lyer illusion. We examined right-brain-damaged patients with ($n = 7$) and without ($n = 7$) left USN, and neurologically unimpaired participants ($n = 14$) in the bisection of Judd stimuli under visual, haptic, and visuo-haptic presentation. Neglect patients showed the characteristic rightward bias in the bisection of the baseline stimuli in the visual modality, but not in the haptic and visuo-haptic conditions. The illusory effects were preserved in each group and in each modality, indicating that the processing of the cross modal illusion is independent of the presence of deficits of spatial attention and representation. Spatial neglect can be modality-specific, but visual and tactile sensory inputs are properly integrated.

**Keywords: unilateral spatial neglect, multisensory integration, illusion, vision, haptics, touch, somatosensory processing**

## INTRODUCTION

Unilateral spatial neglect (USN) is a frequent neuropsychological syndrome occurring after lesions to the right hemisphere. USN is characterized by the patients' failure to report sensory events taking place in the portion of space contralateral to the side of the lesion (contralesional), and to explore through motor acts that portion of space (Vallar, 1998; Halligan et al., 2003; Heilman et al., 2003; Husain, 2008). A rightward bias in line bisection is considered one of the signatures of USN (Bisiach et al., 1976, 1983; Schenkenberg et al., 1980), which has been mainly investigated in the visual modality, with fewer studies assessing the haptic modality, and reporting controversial results.

The aim of this study was twofold: (i) to compare the severity of left USN, as assessed by a line bisection task, in two unimodal visual, and haptic conditions, which were entirely comparable except for the availability of unisensory information; (ii) to assess the patients' ability to combine information from different sensory modalities, i.e., vision and haptics. To this aim, we used a line bisection task involving the processing of a cross modal illusion that we had previously used with neurologically unimpaired individuals (Mancini et al., 2010).

As established by motor exploratory tasks, USN may occur in both the visual and the tactile modality (De Renzi et al., 1970; Beschin et al., 1996; Haeske-Dewick et al., 1996). Evidence has however been provided to the effect that USN may be less severe and even absent in the tactile modality, in the absence of visual input. An early and seminal observation is provided by Gilliatt and Pratt (1952) about a right-brain-damaged (RBD) patient who showed severe left USN, when required to circle pins using a pencil with eyes open; conversely, with eyes closed, the patient explored the whole board up to the extreme left. A number of subsequent studies found a visuo-haptic difference, with USN being more severe in the visual modality for spatial exploratory tasks (Chedru, 1976, in RBD patients with a visual-half-field deficit; Villardita, 1987; Gentilini et al., 1989; Schindler et al., 2006). However, in the study by Chedru (1976) RBD patients without visual half-field deficits showed a more severe USN when blindfolded. Importantly, there is evidence that the deficit may be modality-specific (Cubelli et al., 1991 re-analysis of the data of Gentilini et al., 1989; Vallar et al., 1991b). Particularly, in their re-analysis Cubelli et al. (1991) reported four RBD patients who showed a disproportionate rightward bias with open eyes, but not with eyes closed, in a task requiring to explore a keyboard; three patients showed the opposite pattern (rightward bias with eyes closed), while five patients were impaired in both conditions. Other studies found a double dissociation between visual and tactile USN, reporting patients with a defective performance either in the visual or in the tactile modality (Perani et al., 1987, Appendix 2; Barbieri and De Renzi, 1989; Vallar et al., 1991b).

In line bisection tasks, the available studies indicate that the rightward bias appears to be confined to the visual modality (Fujii et al., 1991; Hjaltason et al., 1993; Chokron et al., 2002). In haptic bisection, no rightward bias has been found, with left USN being almost absent (for a review, see Brozzoli et al., 2006; Gainotti, 2010). Taken together, these findings indicate that USN may be more severe in the visual than in the tactile modality. The deficits may also be modality-specific in exploratory tasks, conjuring up a double dissociation (Vallar, 2000) between vision and touch, and suggesting the existence of modality-specific attentional and representational components (Vallar, 1998). In line with this idea, a rehabilitation study showed that a 6-weeks visual attention training improved visual but not tactile detection of left-sided targets (Làdavas et al., 1994). Finally, it is also well-known that visual and tactile extinction to double simultaneous stimulation may occur independently of each other after unilateral brain damage (Vallar et al., 1994a; Hillis et al., 2006).

The evidence for modality-specificity, currently framed in the broader context of a multi-componential attentional/representational account of the USN deficit (Barbieri and De Renzi, 1989; Vallar, 1998), may be contrasted with an early interpretation of the syndrome in terms of a higher-order sensory impairment, hypothesized as a defective "spatial summation" and termed "amorphosynthesis" (with the primary function of "morphosynthesis" being the recognition of form, Denny-Brown et al., 1952). "*The loss of visual components of such morphosynthesis in addition to tactile factors, is the basis of unawareness of part of extrapersonal space and unawareness of self, without disorder of the concept of space or of body schema. Such unawareness differs considerably from simple loss of sensation*" (ibidem, p. 470). While the very concept of amorphosynthesis is admittedly vague, it appears to suggest the impairment of a higher-order spatial factor, related to different sensory modalities.

In our experimental task we used a well-known visual illusion, the Judd figure, that has proven to be cross modal (Mancini et al., 2011). Visual illusions are a tool for investigating implicit processing in USN, since illusory effects arising from the left side of space can be preserved and do not require perceptual awareness to occur (see Vallar and Daini, 2006, for a review). Among these, the Müller-Lyer illusion consists in a line at which ends two outward- or two inward-oriented arrowheads are located. In line bisection tasks, the Judd variant of this illusion (namely, a line with two identical arrowheads at its ends; Judd, 1899; Holding, 1970) has been used (Ro and Rafal, 1996; Mancini et al., 2010). The bisection of the shaft is shifted toward the tail end under visual and haptic presentation in neurologically unimpaired participants; furthermore, the visual illusion transfers cross modally, modulating bisection in the tactile modality. The spatial correspondence between the visual and the tactile stimuli constituting the cross modal illusion has a crucial role in the cross modal transfer (Mancini et al., 2010).

Although the Müller-Lyer illusion has proven to be useful for studying implicit processing in the contralesional space in patients with left USN, these investigations have been confined to the visual modality: haptic and cross modal illusory effects have not been investigated so far (Vallar and Daini, 2006). We therefore examined right brain-damaged patients with and without left USN in the bisection of the Judd variant of the Müller-Lyer illusion under visual, haptic, and visuo-haptic presentation. We aimed at assessing whether visuo-tactile interactions were preserved in these patients. The ability of RBD patients with left USN to combine visuo-haptic information could also provide an experimental assessment of the "amorphosynthesis" hypothesis.

## MATERIALS AND METHODS

### PARTICIPANTS

Fourteen RBD patients and 14 neurologically unimpaired participants took part in the study (see **Table 1**). RBD patients were subdivided into two groups, with and without left USN (N+/N−), as assessed by a standard neuropsychological battery (**Table 2**). Even if a perusal of **Table 1** suggests that N+ patients may be older (as previously found in larger series of patients, Leibovitch et al., 1998; Gottesman et al., 2008), the age of the participants of the three groups was comparable, as assessed by a one-way analysis of variance [$F(2,25) = 3.16$, $p > 0.60$]. Each participant gave written informed consent to take part in the experiment, which had been approved by the Ethics Committee of the IRCSS Istituto Auxologico Italiano.

The patients' demographic and neurological data are summarized in **Table 1**. All 14 RBD patients had unilateral stroke lesions in the right hemisphere. All patients were right-handed, and had no history or neurological evidence of previous neurological diseases, psychiatric disorders, or dementia. All patients had a normal or corrected-to-normal vision. Contralesional motor, somatosensory, and visual half-field deficits, including extinction to tactile and visual double simultaneous stimuli, were assessed by a standard neurological exam (Bisiach and Faglioni, 1974). Visual field defects were also assessed by kinetic Goldmann perimetry, and by a computerized program testing six different positions in both the left and right hemi-fields, at different eccentricities (3°, 6°, 12°). The visual field of two N− patients (P10 and P14) was tested with our customized program only.

The lesion site and size were assessed by CT or MRI scan. Lesions were mapped for each RBD patient using the MRIcro software (Rorden and Brett, 2000) and were drawn manually onto selected horizontal slices of a standard template brain. **Figure 1** shows the overlapped lesion maps of the 14 RBD patients, separately for patients with and without left USN. In N+ patients the maximum overlap involved the right putamen, insula, and frontal inferior orbital cortices (seven patients); in N− patients the maximum overlap was observed over the right rolandic operculum, the superior temporal pole, and the insula (three patients). Overall, lesions were more extensive in the N+ group (mean volume of the lesion = 126 cc, SD ± 79.51, range 74.44–282.76 cc) than in the N− group (mean volume of the lesion = 41.04 cc, SD ± 54.24, range 1.22–129.16 cc), in line with previously reported evidence in large series of patients (Hier et al., 1983a,b; Leibovitch et al., 1998).

#### *Baseline neuropsychological assessment*

The diagnostic battery assessing the presence of left USN included three visuomotor exploratory tasks (line, letter, and bell cancelation), a reading task, a line bisection task, two copying tasks, and one drawing from memory task (**Table 2**). Patients used their right unaffected hand to perform the tasks. In all tasks, the center of the sheet was aligned with the mid-sagittal plane of the patient's trunk. The Mini Mental State Examination (MMSE) was also given (Grigoletto et al., 1999).

i. *Line bisection.* The patients' task was to mark with a pencil the mid-point of six horizontal black lines (two 10 cm, two

**Table 1 | Demographical and neurological data of 14 right-brain-damaged patients with (N+) and without (N−) USN, and of 14 control (C) neurologically unimpaired participants.**

| Patient | Sex | Age (years) | Education (years) | Months post-onset | Etiology | Neurological deficit | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | M | SS | V |
| **N+** | | | | | | | | |
| 1 | M | 63 | 17 | 16 | I | + | + | + |
| 2 | M | 77 | 17 | 23 | I | + | e | e |
| 3 | F | 83 | 13 | 12 | I | + | + | + |
| 4 | F | 72 | 7 | 22 | I | + | − | − |
| 5 | M | 70 | 17 | 1.5 | I | + | + | + |
| 6 | M | 66 | 5 | 14 | H | + | + | + |
| 7 | M | 71 | 17 | 4 | I | + | + | e |
| Mean | | 71.71 | 13.29 | 13.21 | | | | |
| (SD) | | (6.68) | (5.23) | (8.21) | | | | |
| **N−** | | | | | | | | |
| 8 | F | 41 | 10 | 2 | H | + | − | − |
| 9 | M | 63 | 17 | 15 | I | + | − | − |
| 10 | M | 38 | 13 | 1.5 | I | + | − | − |
| 11 | F | 77 | 8 | 10 | I | − | − | − |
| 12 | M | 74 | 12 | 24 | I | + | − | − |
| 13 | M | 37 | 13 | 1 | I | + | − | − |
| 14 | M | 39 | 6 | 1 | I | + | − | − |
| Mean | | 52.71 | 11.29 | 7.79 | | | | |
| (SD) | | (17.96) | (3.64) | (8.99) | | | | |
| **C** | | | | | | | | |
| 15 | F | 72 | 8 | | | | | |
| 16 | M | 58 | 17 | | | | | |
| 17 | F | 60 | 8 | | | | | |
| 18 | M | 52 | 13 | | | | | |
| 19 | M | 52 | 8 | | | | | |
| 20 | M | 65 | 5 | | | | | |
| 21 | F | 53 | 13 | | | | | |
| 22 | M | 66 | 13 | | | | | |
| 23 | M | 73 | 17 | | | | | |
| 24 | M | 85 | 16 | | | | | |
| 25 | F | 70 | 13 | | | | | |
| 26 | F | 85 | 13 | | | | | |
| 27 | F | 36 | 8 | | | | | |
| 28 | F | 41 | 13 | | | | | |
| Mean | | 62.00 | 11.79 | | | | | |
| (SD) | | (14.64) | (3.77) | | | | | |

*M/SS/V: left motor/somatosensory/visual half-field deficits. e: Contralesional extinction.*
*±Presence/absence of impairment.*
*M/F, male/female; I/H, ischemic/hemorrhagic.*

15 cm, and two 25 cm in length, all 2 mm in width), presented in a random fixed order. Each line was printed at the center of an A4 sheet. The length of the left-hand side of the line (i.e., from the left end of the line to the participant's mark) was measured to the nearest millimeter. This measure was converted into a standardized score (percentage deviation), namely: (measured left half minus objective half)/objective half × 100 (Rode et al., 2006b). This transformation yields positive numbers for marks placed to the right of the objective physical center, negative numbers for marks placed to the left of it. The mean percentage deviation score of 65 neurologically unimpaired participants, matched for age (mean = 72.2, SD ± 5.16, range 65–83), and years of education (mean = 9.5, SD ± 4.48, range 5–18) was 1.21% (SD ± 3.48, range −16.2% to +6.2%). A percentage deviation score higher than 8.20 was considered as indicative of left USN (Fortis et al., 2010).

ii. *Line cancellation* (Albert, 1973). The participants' task was to cross out all of the 21 black lines printed on an A4 sheet

**Table 2 | Clinical data of 14 right-brain-damaged patients with (N+) and without (N−) USN.**

| Patient | MMSE | Cancellation tests | | | Line bisection (%) | Sentence reading | Drawing | | | Corsi's block tapping test | | Tactile form recognition |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Line | Letter | Bells | | | Daisy | Complex | Clock | Standard | Vertical | |
| **N+** | | | | | | | | | | | | |
| 1 | 30 | 0 | 2 | 1 | +11.4* | 6 | 2 | 5.5* | 11 | 2.75* | 3 | 30 |
| 2 | 29 | 0 | 20* | 16* | +3.8 | 6 | 1.5 | 10 | 12 | 3.75 | 2 | 25 |
| 3 | 26 | 2* | 6* | 10* | +3.8 | 4* | 1.5 | 4* | 2* | 3.5 | 1* | 10* |
| 4 | 27 | 2* | −1 | 3 | +10.2* | 6 | 2 | 9* | 11 | 3.5 | 2 | 20* |
| 5 | 29 | 0 | 36* | 7* | +1.1 | 6 | 1.5 | 4.5* | 6* | 1.75* | 1* | 16* |
| 6 | 28 | 7* | 34* | 6* | +50.4* | 0* | 1.5 | 4* | 8* | 2.25* | 2 | 14* |
| 7 | 28 | 1* | 32* | 7* | +2.4 | 4* | 2 | 5* | 12 | 3.75 | 2 | 25 |
| **N−** | | | | | | | | | | | | |
| 8 | 30 | 0 | 0 | 0 | +4.4 | 6 | 2 | 10 | 12 | 3.75 | 2 | 28 |
| 9 | 29 | 0 | 0 | −2 | −2.3 | 6 | 2 | 10 | 12 | 5.75 | 2 | 33 |
| 10 | 30 | 0 | 0 | 0 | −2 | 6 | 2 | 10 | 12 | 3.5 | 3 | 29 |
| 11 | 27 | 0 | 0 | 0 | −5.6 | 6 | 2 | 10 | 12 | 4.5 | 3 | 24 |
| 12 | 27 | 0 | −1 | −2 | −5.6 | 6 | 2 | 10 | 12 | 4 | 3 | 22* |
| 13 | 28 | 0 | 0 | 1 | +0.8 | 6 | 2 | 10 | 12 | 3.5 | 4 | 36 |
| 14 | n.a.[1] | 0 | 1 | 4 | +0.2 | 6 | 2 | 10 | 11 | 4 | 4 | 34 |

*Target cancellation: numbers of left- minus right-hand-side omissions. Line bisection: percent displacement (±rightward/leftward). Sentence reading, drawing tests, Corsi's block tapping tests, tactile form recognition: number of correct responses (for Corsi's standard tapping test, adjusted scores). n.a.: not assessed. [1]Arabian patient. *defective performance.*

with no distracters. The score was the difference between numbers of omissions in the left- (range 0–11) and in the right- (range 0–10) hand-sides of the sheet. Neurologically unimpaired participants perform this task without errors.

iii. *Letter cancellation* (Diller and Weinberg, 1977). The participants' task was to cross out all of 104 H letters (53 in the left-hand-side, and 51 in the right-hand-side of the sheet), printed on an A3 sheet, together with 208 letter distracters. In neurologically unimpaired participants the maximum difference between omission errors on the two sides of the sheet is two (Vallar et al., 1994b).

iv. *The Bells test* (Gauthier et al., 1989). The participants' task was to cross out all of 35 bells (15 in the left-hand-side, 5 in the middle, and 15 in the right-hand-side of the sheet), printed on an A3 sheet, together with other 280 distracters. In neurologically unimpaired participants the maximum difference between omission errors on the two sides of the sheet is four (Vallar et al., 1994b).

v. *Sentence reading* (Pizzamiglio et al., 1992). Six sentences of different lengths were presented one at time, printed centrally on a A4 sheet. The score was the number of correctly read sentences (range 0–6). Normal participants and patients with right brain damage without USN make no errors on this test. RBD patients with USN make omission errors, substitution errors, or both, in the left half of the sentence.

vi. *Drawing.* Patients were required to copy two figures [a daisy and a complex figure with two trees in the left-hand-side, two pine trees in the right-hand-side, and a house in the center of an A4 sheet (Gainotti et al., 1972)], and to draw from memory the hours of a clock in a circular quadrant (diameter 12 cm), printed on an A4 sheet. Omission errors were calculated as follows:

   a. *Daisy* (range 0–2): 2 (flawless copy); 1.5 (partial omission of the left-hand-side of the daisy); 1.0 (complete omission of the left-hand-side of the daisy); 0.5 (complete omission of the left-hand-side of the daisy, and partial omission of the right-hand-side of the daisy); 0 (no drawing, or no recognizable element). The mean omission score of 148 neurologically unimpaired participants (mean age = 61.89, SD ± 11.95, range 40–89) was 1.99 (SD ± 0.12, range 1–2). Accordingly, the presence of a partial or complete omission of the left-hand side of the daisy (score lower than 1.5) was considered as indicative of left USN.

   b. *Five-element complex drawing* (range 0–10): 2 (for each flawless copied element); 1.5 (for each partial left-sided omission of one component, e.g., some branches of the left-hand-side of a tree); 1.0 (for each complete left-hand-side omission of one component); 0.5 (for each complete omission of the left-hand-side, and partial omission of the right-hand-side of the component); 0 (no drawing, or no recognizable element). The horizontal ground line was not considered for scoring. The mean score of 148 neurologically unimpaired participants (mean age = 61.89, SD ± 11.95, range 40–89) was 9.89 (SD ± 0.23, range 9.5–10). Accordingly, a score lower than 9.5 indicated a defective performance.

   c. *Clock drawing from memory* (range 0–12): 1 (for each element in the correct position); 0 (for each omission or translocation of an element from one side to the

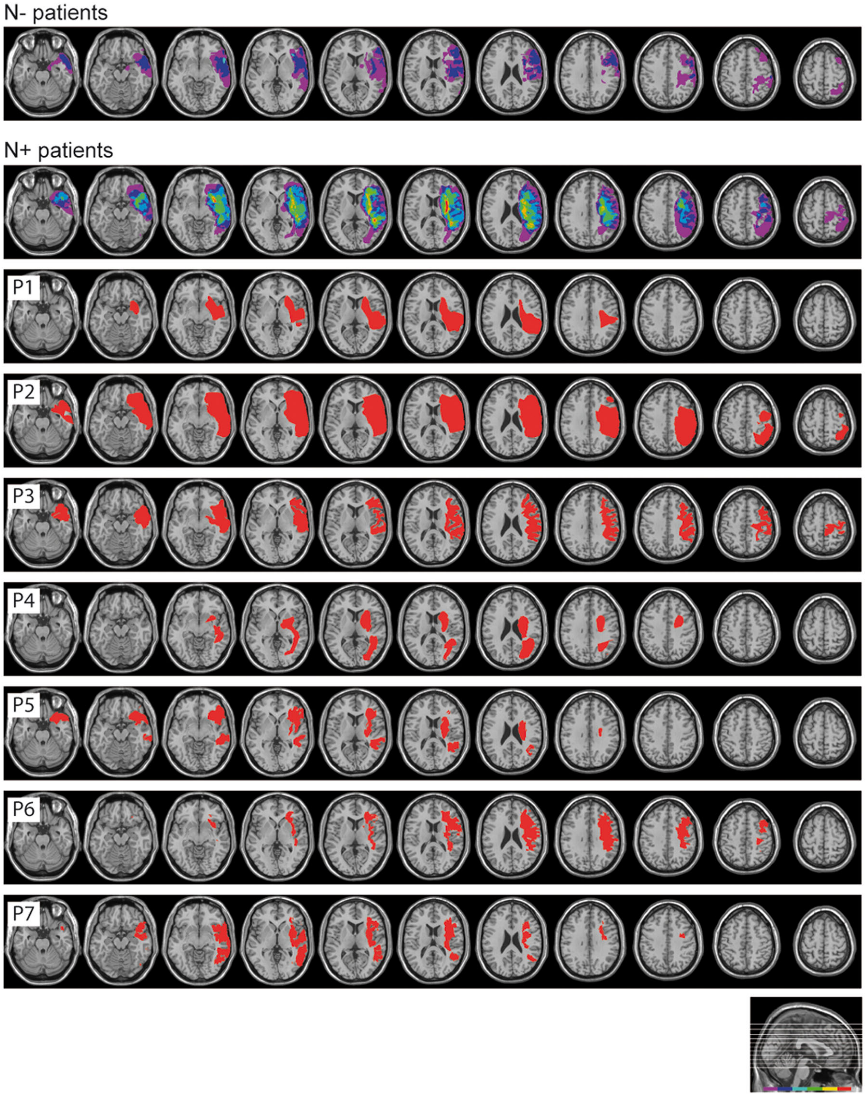

**FIGURE 1 | Lesion site of patients with and without USN.** Superimposed overlapping brain lesions (first two rows), and lesional mapping for each USN patient (P1–P7). The lesions were mapped using MRIcro software (http://www.sph.sc.edu/comd/rorden/mricro.html).

other; elements "12" and "6" were scored as translocated when displaced in the right- or left-hand-side quadrants). The mean score of 148 neurologically unimpaired participants (mean age = 61.89, SD ± 11.95, range 40–89) was 11.55 (SD ± 1.17, range 0–6). Accordingly, a score lower than 9 indicated a defective performance. Furthermore, neurologically unimpaired participants made no translocations.

vii. *Corsi's block tapping test* (Orsini et al., 1987). Nine white cubes were arranged over a 23 by 28 cm board. The examiner tapped sequences of increasing length in a fixed order, with the patients' task being to tap the same ordered sequence, immediately after presentation. The test continued until the patient failed at a given length (less than three out of five sequences were correctly recalled). The spatial span score was the length of the longest sequence correctly recalled. Scores adjusted for gender, age, and education were computed.

viii. *Corsi's block tapping vertical test.* This was a modified version of the standard Corsi's block tapping test, adapted for USN patients (Ronchi et al., 2009). Nine white cubes were arranged over a vertical board 60 cm high and 14 cm wide; the distance between each cube was 1.5 cm. The procedure was identical to that used for the standard block tapping test of Orsini et al. (1987). Control data were provided by 14 neurologically unimpaired right-handed C participants, matched for age and education (mean age 62 years, range 36–85, mean education 11.7 years): the mean span was 3.38 (SD ± 1.07, range 2–6).

ix. *Modified version of the Benton tactile form assessment* (Benton, 1994). We adapted the original version in order to administer the test to USN patients (symmetrical stimuli, central presentation of the visual comparisons). Participants were required to match a shape, explored haptically with the right hand and out of sight, to a visual sketch of the shape to be chosen among five stimuli, printed in a vertical column on an A4 sheet. Eighteen shapes, subdivided in three sections, were presented: nine filled and nine unfilled. A score of 2 was assigned to each correct response given within 30 s, 1 within 60 s, and 0 for wrong or out-of-time responses. The mean score of the control group (C) was 30.91 out of 36 (SD ± 4.18, range 26–36).

For the Benton tactile form assessment and the Corsi's block tapping vertical test, the patients' performances were compared with those of control participants by *t* tests (Crawford and Garthwaite, 2002).

### STIMULI AND APPARATUS

Stimuli and apparatus were identical to a previous study we conducted in neurologically unimpaired participants (Mancini et al., 2010). Stimuli consisted of three types of black plastic figures (**Figures 2A–C**): a baseline control (a horizontal rod with vertical ends), and two illusory figures (a horizontal rod with leftward outgoing/rightward ingoing fins at its ends, which brought about a leftward displacement of the shaft's perceived center; a horizontal rod with leftward ingoing/rightward outgoing fins, which brought about a rightward displacement of the shaft's perceived center). Each stimulus included a horizontal rod (10 or 12 cm long), and two identical ends, vertical (length: 25 mm; height: 10 mm; thickness: 1 mm), or angled at 45° (length of each fin: 35 mm; height: 10 mm; thickness: 1 mm). All stimuli, both with vertical and angled ends, were 50 mm high. Each stimulus configuration was attached in the center of a white wooden board (40 cm × 40 cm, thickness 0.8 cm). Under visual and haptic presentation, both the arrowheads and the horizontal rod were glued on the front of the board (see **Figures 2A,B**). Conversely, in the cross modal condition the horizontal rod was positioned on the backside of the board centrally, and the arrowheads on the front-side in the correspondent positions (see **Figure 2C**).

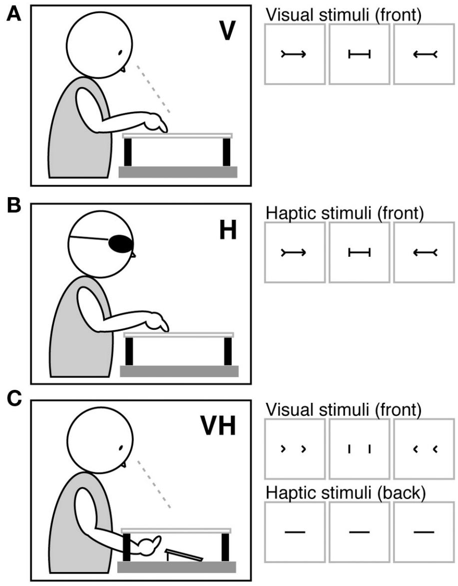

**FIGURE 2 | Stimuli and apparatus.** Under visual [V **(A)**], haptic [H **(B)**], and cross modal visuo-haptic [VH **(C)**] presentations, three types of stimuli were administered: leftward outgoing/rightward ingoing fins, which brought about a leftward displacement of the shaft's perceived center; a baseline control stimulus with vertical ends; leftward ingoing/rightward outgoing fins, which brought about a rightward displacement of the shaft's perceived center. In the visuo-haptic condition **(C)** the ends were glued on the front of the board, and the horizontal shaft to be bisected on the back, in the correspondent positions.

The experiment was performed in a normally illuminated and quiet room with patients being comfortably seated in front of a table. Each board was presented individually, with its center aligned with the mid-sagittal plane of the participant's trunk, and placed flat on a wooden support at the height of 12 cm from the tabletop. In the cross modal condition, a mirror was placed on the table under the board. The mirror reflected the shaft on the backside and was seen by the experimenter only.

### PROCEDURE

Participants received instructions to bisect with the index finger the horizontal shaft of each stimulus, using their right hand. The task was performed under visual, haptic, or visuo-haptic conditions of stimulus presentation. The three conditions were assessed during three separate sessions, in a counterbalanced order across participants. In the *visual condition,* participants received instructions to touch the mid-point of the shaft without exploring haptically the stimulus, and to close their eyes immediately after responding, while the experimenter measured their bisection error to the nearest millimeter (**Figure 2A**). In the *haptic condition*, blindfolded participants scanned the shapes (arrowheads and shafts) haptically, and then set the mid-point of the shaft with their right index finger. Each trial started with the experimenter placing the palm of the participant's open hand centrally over the stimulus. Stimuli were short enough to fit into the open hand. On each trial, participants were required to explore the entire shape before responding, with no time limits being set (**Figure 2B**). In the *visuo-haptic condition*, participants received instructions to look at the arrowheads on the front-side of the board, and simultaneously palpated with the whole hand the shaft glued on the backside of the board, and then set its mid-point using their right index finger. Participants did not see their right forearm, which was covered by the wooden support (**Figure 2C**).

In all presentation conditions (visual, haptic, visuo-haptic) the two lengths of the shaft (10, 12 cm), and the three types of stimulus configuration (baseline with vertical ends; leftward outgoing/rightward ingoing fins; leftward ingoing/rightward outgoing fins) generated six possible stimuli. Each type of stimulus was presented eight times, for a total of 48 trials. For each participant and for each session, a different random sequence was used. Two practice trials, one baseline, and one illusory stimulus selected at random, were administered at the beginning of each session (visual, haptic, visuo-haptic), and were not included in any subsequent analyses.

After each response, using a ruler, the experimenter measured to the nearest millimeter the bisection error, namely, the distance between the subjective mid-point set by each participant and the objective center of the shaft. A rightward deviation of the subjective mid-point carried a plus sign, a leftward deviation a minus sign. A percent error as related to the length of each shaft (10, 12 cm) was subsequently computed.

In order to assess the presence of USN in each sensory modality, percent errors of the baseline bisection task were first submitted to three analyses of variance (ANOVA), one per Presentation Modality, with one within-subjects main factor (*Shaft Length*: 10, 12 cm), and one between-subjects main factor (*Group*: N+, N−, C).

Subsequently, for each participant, average percent errors in each illusory stimulus condition were corrected for the average percent error in the baseline stimulus. Illusory effects were investigated by three separate ANOVAs (one per Presentation Modality) with two within-subjects main factors (*Stimulus:* leftward outgoing/rightward ingoing fins, leftward ingoing/rightward outgoing fins; *Shaft Length*: 10, 12 cm), and one between-subjects main factor (*Group*: N+, N−, C). *A posteriori* contrasts among means were evaluated by Scheffé's test.

## RESULTS

### BASELINE ERRORS

The bisection of the baseline stimulus (a shaft with vertical ends) was deviated rightward in the N+ group in the visual modality, indicating the presence of visual USN. No difference across the three groups was found under unimodal haptic and cross modal visuo-haptic presentations (see **Figure 3**).

In the *visual condition*, the analysis of variance revealed a significant main effect of Group [$F(2,25) = 20.32$, $p < 0.0001$, $\eta^2 = 0.399$]. The main effect of Shaft Length was significant [$F(1,25) = 5.31$, $p = 0.030$, $\eta^2 = 0.161$], indicating larger rightward errors with 12 cm than with 10 cm stimuli. The interaction Shaft Length by Group was not significant [$F(2,25) = 1.32$, $p = 0.285$, $\eta^2 = 0.080$]. The differences between N+ patients and both N− ($p < 0.0001$) and C participants ($p < 0.0001$) were significant; no difference was found between the N− and C groups.

In the *haptic condition*, the main effect of Group was not significant ($F < 1$). The main effect of Shaft Length was significant [$F(1,25) = 6.09$, $p = 0.021$, $\eta^2 = 0.151$], since the 10 cm stimulus was bisected more leftward than the 12 cm one. The interaction Shaft Length by Group was significant [$F(2,25) = 4.55$, $p = 0.021$, $\eta^2 = 0.226$]. The difference between short and long stimuli was significant only in the N+ group ($p = 0.001$): in particular, N+ patients bisected the longer rod more rightward than the shorter one.

In the *visuo-haptic condition*, the main effect of Group was not significant ($F < 1$). Also the main effect of Shaft Length [$F(1,25) = 1.13$, $p = 0.298$, $\eta^2 = 0.042$] and its interaction with Group ($F < 1$) were not significant.

### ILLUSORY EFFECTS

In each modality and in each group, stimuli with leftward outgoing/rightward ingoing fins brought about a leftward error, stimuli with leftward ingoing/rightward outgoing fins elicited a rightward error (**Figure 4**). The figure does not show the effect of Shaft Length, which did not provide results of interest for the purposes of the present study.

In the *visual condition*, the analysis of variance did not reveal a significant main effect of Group [$F(2,25) = 2.51$, $p = 0.101$, $\eta^2 = 0.148$]; the main effect of Stimulus [$F(1,25) = 138.88$, $p < 0.0001$, $\eta^2 = 0.772$], and its interaction with the main effect of Shaft Length [$F(1,25) = 6.60$, $p = 0.017$, $\eta^2 = 0.004$] were significant. *Post hoc* comparisons did not show any significant difference between the two lengths for both the leftward outgoing/rightward ingoing ($p = 0.322$), and the leftward ingoing/rightward outgoing stimuli ($p = 0.236$). The main effect of Shaft Length and all the interactions with Group were not significant ($F < 1$).

In the *haptic condition*, there was a trend toward significance for the main effect of Group [$F(2,25) = 3.11$, $p = 0.062$, $\eta^2 = 0.182$]. N− patients bisected the stimuli overall more rightward than both the N+ ($p = 0.035$) and the C ($p = 0.037$) participants. The main effect of Stimulus [$F(1,25) = 45.663$, $p < 0.0001$, $\eta^2 = 0.501$] and its interaction with the main effect of Shaft Length [$F(1,25) = 4.96$, $p = 0.035$, $\eta^2 = 0.006$] were significant.

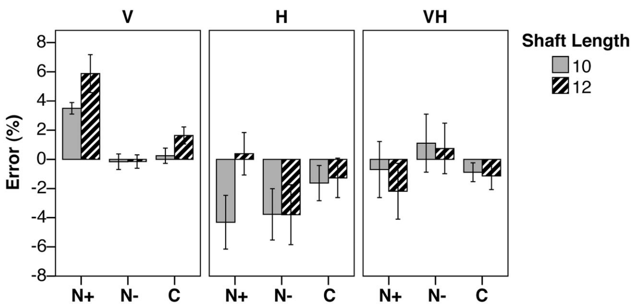

**FIGURE 3 | Baseline errors.** Mean percent error (±SEM) in shaft bisection in the baseline condition, by Shaft Length (10, 12 cm), Group (N+/N−, patients with/without USN; C, neurologically unimpaired control participants), and presentation Modality (Visual, V; Haptic, H, Visuo-Haptic, VH). Negative/positive score: leftward/rightward error.

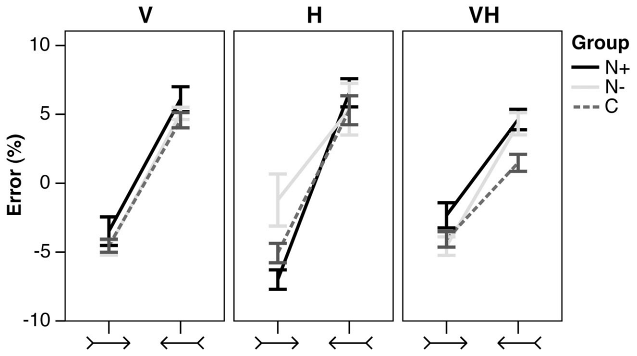

**FIGURE 4 | Illusory effects.** Mean percent error (±SEM) in shaft bisection in the illusory conditions, by Stimulus type (leftward outgoing/rightward ingoing, and leftward ingoing/rightward outgoing fins), Group (N+, N−, C), and presentation Modality (V, H, VH). Negative/positive score: leftward/rightward error.

The difference between the two lengths was close to significance in the leftward ingoing/rightward outgoing Stimulus ($p = 0.076$). The main effect of Shaft Length and its interaction with the main effect of Group was not significant ($F < 1$). The Stimulus by Group [$F(2,25) = 1.50, p = 0.243, \eta^2 = 0.033$], and Stimulus by Length by Group [$F(2,25) = 1.43$, $p = 0.258$, $\eta^2 = 0.003$] interactions were not significant.

Finally, in the *visuo-haptic condition*, the main effect of Group was significant [$F(2,25) = 5.92, p = 0.008, \eta^2 = 0.321$]. The difference between the N+ and the C groups was significant ($p < 0.009$): the corrected bisection of the illusory stimuli made by N+ patients was shifted overall more rightward than the bisection made by C participants. No other difference was significant. The main effect of Stimulus [$F(1,25) = 113.01, p < 0.0001, \eta^2 = 0.622$] was significant. The main effect of Shaft Length ($F < 1$), and the Stimulus by Group [$F(2,25) = 2.18$, $p = 0.123$, $\eta^2 = 0.025$], Stimulus by Shaft Length [$F(1,25) = 1.68$, $p = 0.206$, $\eta^2 = 0.004$], Shaft Length by Group ($F < 1$), and Stimulus by Shaft Length by Group [$F(2,25) = 1.36$, $p = 0.274$, $\eta^2 = 0.007$] interactions were not significant.

## DISCUSSION

The present study yielded three main findings. (1) The rightward bias in bisection characteristic of patients with USN was present only in the visual modality, not in the haptic and visuo-haptic conditions. (2) Illusory effects were preserved in each modality, revealing that the processing of the illusion is independent of the presence of spatial attentional deficits. (3) Visual and tactile inputs were properly integrated by RBD patients with left USN.

### USN WITHIN AND BETWEEN SENSORY MODALITIES

Our results support the view that left USN can be modality-specific. In the present study, USN (as evaluated by a bisection task) was present in the visual modality only, since USN patients showed a preserved performance in the haptic and cross modal presentation conditions. It may be noted, however, that in the haptic modality USN patients did make a more rightward error with the longer stimulus (**Figure 3**), unlike patients without USN and control participants. This pattern might be taken as possible evidence of a minor tactile USN. In the visuo-haptic condition (visual illusion and haptic line bisection), illusory effects were biased rightwards in USN patients as compared with control participants (**Figure 4**), possibly indicating transfer of the visual USN in the tactile domain.

Overall, these results are in line with the evidence that USN is absent or less severe in the tactile than in the visual modality: a number of studies reports almost preserved tactile bisection in patients with visual USN (Fujii et al., 1991; Hjaltason et al., 1993; Chokron et al., 2002). We cannot exclude that in the present experiment the use of short rods (<12 cm) might have hidden deficits in haptic bisection: particularly, this is suggested by the fact that USN patients show a small rightward bias (see **Figure 3**) in the haptic bisection of longer (12 cm), but not of smaller rods (10 cm). However, longer rods have been used in previous studies that do not report a greater rightward error in haptic line bisection in USN patients than in control participants (Fujii et al., 1991, four lengths from 8 to 20 cm in 4 cm steps; Hjaltason et al., 1993, 20, and 40 cm; Chokron et al., 2002, 20, and 22 cm). Interestingly, the experiment of Hjaltason et al. (1993) includes a visuo-tactile task, in which participants indicate the perceived mid-point after having ran the index finger along the rod. Even in that condition (which differs from the visual one in that tactile exploration of the rod is required), RBD patients with left USN show no significant rightward error. In our visuo-haptic condition, the rod to bisect is explored only haptically, and participants are allowed to see its ends.

It is possible that our selection criteria for USN, based on visual tasks, may have prevented the inclusion of patients with haptic USN, at least in part accounting for the present results in terms of selection bias. However, it should be noted that previous studies reporting the absence of USN in haptic line bisection adopted similar selection criteria, namely the presence of visual USN (Fujii et al., 1991; Hjaltason et al., 1993; Chokron et al., 2002). In any case, in the baseline experimental task we assessed visual and haptic (as well as visuo-haptic) line bisection, replicating the previously reported visuo-haptic dissociation.

The rightward bias exhibited by RBD USN patients in visual bisection can not be traced back to initial rightward biases or to a general position preference for the side ipsilateral to the side of the lesion (ipsilesional; Campbell and Oxbury, 1976; Costa, 1976), since the shafts are short (10 and 12 cm) and the participants' open hand covered the whole stimulus at the beginning of each trial.

Multisensory interactions have been rarely investigated in USN. One area of research involves the effects of physiological stimulations that improve a number of manifestations of the USN syndrome (Vallar et al., 1997; Kerkhoff, 2003; Rode et al., 2006a; Chokron et al., 2007). Vestibular stimulation ameliorates left-sided somatosensory deficits (Vallar et al., 1990, 1993b; Bottini et al., 2005). Prism adaptation improves tactile deficits (Maravita et al., 2003). Optokinetic stimulation may either ameliorate or worsen proprioceptive deficits of position sense in RBD patients with left USN (Vallar et al., 1993a, 1995a). Also auditory USN, as indexed by a contralesional left-sided deficit with dichotic stimuli, is ameliorated by prism adaptation (Jacquin-Courtois et al., 2010). The effects of these stimulations are essentially similar across sensory modalities and depend on the characteristics of the sensory stimulations (e.g., side, left vs. right, and type, warm vs. cold stimulation, in the case of caloric vestibular stimulation: see reviews in Vallar et al., 1997; Kerkhoff, 2003; Rode et al., 2006a; Chokron et al., 2007). Importantly, these stimulations are thought to modulate spatial processing or attention (biased ipsilaterally in USN patients), rather than being considered as an index of preserved multisensory integration. In particular, the typical paradigm of these studies involves the assessment of the patients' performance in unimodal tasks.

More direct evidence comes from the finding that in RBD patients the detection of contralesional visual stimuli is improved by the concomitant presentation of stimuli in another sensory modality (i.e., auditory), depending on the temporal and spatial coincidence of the sensory inputs (Frassinetti et al., 2002b, 2005). Differently from the effects of the sensory stimulations discussed above, these effects have been interpreted as based on the integrative contribution of multisensory neurons and neural networks, spared in patients with USN (Stein and Stanford, 2008). In particular, since these cross modal effects have been found both in patients with sensory deficits (namely, hemianopia) and with visuo-spatial deficits (namely, USN), the underlying mechanisms might involve preserved multisensory integration, with effects similar to those found in neurologically unimpaired participants (Frassinetti et al., 2002a; Bolognini et al., 2005). These studies investigated the multisensory integration of multiple sources of information about the same stimulus, i.e., coincident visual and auditory targets. A different form of integration is the *combination* between non-redundant sensory inputs, necessary to form a robust and coherent representation (Ernst and Bulthoff, 2004). The present study aimed specifically at investigating whether multisensory combination is preserved in USN, using the transfer of the illusion from vision to haptics as an index of effective integration.

The present results suggest that multisensory combinations, over or above contralesional unimodal sensory deficits, are preserved in RBD patients with USN, and are therefore independent of the presence of spatial deficits. This conclusion is further supported by two findings: first, patients exhibit visual but not, or minor, haptic USN; second, illusory effects are not modality-specific. Interestingly, in the visuo-haptic condition (**Figure 4**), the finding that N+ patients show a greater overall rightward bias with respect to control participants may be taken as a further evidence of visuo-haptic integration, as if visual USN cross modally affected tactile bisection. The preserved multisensory interactions reported here do not extend to other haptic processes, such as shape recognition, that was defective in four out of seven N+ patients and in one out of seven N− patients. Overall, shape recognition and cross modal integration appear to rely on largely independent processes, which, in turn, do not involve the spatial attentional resources defective in USN.

Consequently, results from the present visuo-haptic paradigm do not lend support to the "amorphosynthesis" hypothesis of Denny-Brown et al. (1952), even though this account was formulated in rather vague terms. More recently, Brandt et al. (2009) proposed that USN reflects the damage of a multisensory integration center for attention and orientation (MSO) in the temporo-parietal cortex. The MSO is assumed to be bilateral, but the center localized in the right hemisphere is held to be dominant, in that it exerts a greater inhibition of the contralateral left MSO, and a greater excitation of the ipsilateral visual cortex. The net result is that a right-sided temporo-parietal lesion of the MSO brings about visual USN mainly through a reduced activity of the right-sided visual cortex, which is further inhibited by the contralateral visual cortex. This model considers USN mainly as a visual phenomenon, and therefore could seem in accordance with the present results at first analysis. However, USN has been found also in the tactile (Vallar et al., 1991a, 1993b; Smania and Aglioti, 1995) and auditory (Bisiach et al., 1984; Vallar et al., 1995b; Jacquin-Courtois et al., 2010) modalities, in line with the multi-componential nature of the disorder (Vallar, 1998). Also, the multisensory integration features of the MSO center do not appear supported by the present results, which clearly reveal preserved visuo-haptic interactions in RBD patients with left USN.

### PROCESSING OF THE JUDD ILLUSION

The illusory effects are preserved in each sensory condition, and independent of the presence of USN. Preserved leftward illusory effects have been already demonstrated in the visual modality, using variants of the Müller-Lyer figure (see Daini et al., 2002; Vallar and Daini, 2006, for reviews). These findings are in striking contrast with the evidence that the explicit processing of the left-sided portion of such stimuli is defective, as assessed by the verbal report of left-sided fins (Mattingley et al., 1995) and by same-different judgments (Ro and Rafal, 1996; Olk et al., 2001). Here we demonstrate for the first time that also tactile and cross modal illusory effects are preserved in RBD patients with left visual USN. Therefore, the Judd illusion can be a powerful tool for evaluating multimodal visual, haptic, and cross modal processes.

The present findings provide a definite indication that the spatial and attentional resources disrupted by USN do not play an important role in the processing of illusions such as the Judd variant of the Müller-Lyer figure. These processes may be largely non-spatial in nature, with the illusion eliciting a bias in mechanisms involved in cross modal shape representation in ventral stream networks (Vallar and Mancini, 2010; Mancini et al., 2011). A recent study from our laboratory indicates that a region in the extra-striate visual cortex, the lateral occipital complex, is a crucial underpinning of the multisensory Judd illusion (Mancini et al., 2011). This region is not a lesional correlate of USN, as assessed by line bisection tasks. In our patients the cortico-subcortical lesions mainly involve frontal and centro-parietal regions, relatively sparing the extra-striate visual cortex (see **Figure 1**). Anatomo-clinical correlation studies in RBD patients with USN show that the rightward bias in line bisection is associated with posterior lesions (Binder et al., 1992): specifically, to the inferior parietal lobule (Mort et al., 2003; Verdon et al., 2010), and at the temporo-occipital junction (Rorden et al., 2006).

In the present study the illusory effects were assessed by a line bisection task. This method, which provides a measure of the Müller-Lyer and related illusions, has been largely used in neuropsychological investigations of the processing of illusory visual stimuli (review in Vallar and Daini, 2006), and specifically using the Judd illusion, in previous studies from our laboratory (Mancini et al., 2010, 2011). In studies performed in neurologically unimpaired participants other response effectors have been investigated, such as saccadic eye movements, visually guided pointing, and grasping, with different results. Saccades are biased by the illusion, suggesting no dissociation between this type of action and perception (see the meta analysis of Bruno et al., 2010). As for visual effects, the prediction could be made that USN patients, being sensitive to illusory effects as assessed by line bisection, would exhibit a modulation of saccades by illusory stimuli. Beside saccades, the visual illusion can also affect grasping movements (Bruno and Franz, 2009). Conversely, visually guided pointing is not affected by the illusion (see the meta analysis of Bruno et al., 2008). This finding is taken as broadly consistent with the distinction between vision-for-perception and vision-for-action (Milner and Goodale, 2006; Gangopadhyay et al., 2010), with the illusion arising in the visual perceptual ventral stream and not affecting the vision-for-action dorsal stream. However, the positive effects on saccades and grasping "do not appear to support independent spatial representations for vision-for-action and vision-for-perception" (Bruno and Franz, 2009).

## CLINICAL IMPLICATIONS

USN includes both perceptual and premotor components, with the former involving a defective conscious spatial representation of sensory and internally generated events in the contralesional side, the latter an impairment in performing movements in a contralesional direction ("directional hypokinesia"), and a general ipsilesional bias (Vallar and Mancini, 2010). The present study did not aim at disentangling perceptual and premotor components of USN. This would require specific paradigms, which should contrast perception and action in a more or less compatible way. Nevertheless, we believe that the rightward bias we found in the visual bisection of baseline stimuli is likely to be mainly perceptual in nature for three main reasons. (i) The stimuli (10 and 12 cm in length) fit comfortably into the participants' hand, which was placed over the stimulus at the beginning of the trial, thus minimizing the need of manual exploration. (ii) The preserved illusory effects, as assessed by manual line bisection, involve both rightward and leftward shifts, performed by the unaffected right hand. (iii) Premotor pathological mechanisms appear to be less frequent determinants of USN than the perceptual deranged components (Gallace et al., 2008; Vallar and Mancini, 2010).

Finally, on a clinical note, the present findings that USN can be absent in the tactile domain and spare cross modal interactions support the importance of including a multimodal assessment in diagnostic batteries, and of setting up multisensory-based rehabilitation approaches rather than the traditional visual treatments (Pizzamiglio et al., 2006; Schroder et al., 2008). The most important functions of multisensory integration are likely to be maximizing information delivered from the different sensory modalities, reducing the variance in the multisensory sensory estimate, in order to increase its reliability (Ernst and Bulthoff, 2004). Left USN may cause a bias in one modality, but the brain can take advantage of other preserved sensory modalities to help correcting it. Treatments that support these processes should be encouraged.

## ACKNOWLEDGMENTS

This work has been supported in part by University of Milano-Bicocca FAR grants to Emanuela Bricolo and Giuseppe Vallar, and by Ricerca Corrente grants to Giuseppe Vallar, from the Italian Auxological Institute, Milan, Italy. Flavia Mancini has been supported by a doctoral program scholarship of the University of Milano-Bicocca, Italy.

## REFERENCES

Albert, M. L. (1973). A simple test of visual neglect. *Neurology* 23, 658–664.

Barbieri, C., and De Renzi, E. (1989). Patterns of neglect dissociation. *Behav. Neurol.* 2, 13–24.

Benton, A. L. (1994). Neuropsychological assessment. *Annu. Rev. Psychol.* 45, 1–23.

Beschin, N., Cazzani, M., Cubelli, R., Della Sala, S., and Spinazzola, L. (1996). Ignoring left and far: an investigation of tactile neglect. *Neuropsychologia* 34, 41–49.

Binder, J., Marshall, R., Lazar, R., Benjamin, J., and Mohr, J. P. (1992). Distinct syndromes of hemineglect. *Arch. Neurol.* 49, 1187–1194.

Bisiach, E., Bulgarelli, C., Sterzi, R., and Vallar, G. (1983). Line bisection and cognitive plasticity of unilateral neglect of space. *Brain Cogn.* 2, 32–38.

Bisiach, E., Capitani, E., Colombo, A., and Spinnler, H. (1976). Halving a horizontal segment: a study on hemisphere-damaged patients with cerebral focal lesions. *Schweiz. Arch. Neurol. Psychiatr.* 118, 199–206.

Bisiach, E., Cornacchia, L., Sterzi, R., and Vallar, G. (1984). Disorders of perceived auditory lateralization after lesions of the right hemisphere. *Brain* 107, 37–52.

Bisiach, E., and Faglioni, P. (1974). Recognition of random shapes by patients with unilateral lesions as a function of complexity, association value and delay. *Cortex* 10, 101–110.

Bolognini, N., Frassinetti, F., Serino, A., and Làdavas, E. (2005). "Acoustical vision" of below threshold stimuli: interaction among spatially converging audiovisual inputs. *Exp. Brain Res.* 160, 273–282.

Bottini, G., Paulesu, E., Gandola, M., Loffredo, S., Scarpa, P., Sterzi, R., Santilli, I., Defanti, C. A., Scialfa, G., Fazio, F., and Vallar, G. (2005). Left caloric vestibular stimulation ameliorates right hemianesthesia. *Neurology* 65, 1278–1283.

Brandt, T., Glasauer, S., Strupp, M., and Dieterich, M. (2009). Spatial neglect: hypothetical mechanisms of disturbed interhemispheric crosstalk for orientation. *Ann. N. Y. Acad. Sci.* 1164, 216–221.

Brozzoli, C., Demattè, M. L., Pavani, F., Frassinetti, F., and Farnè, A. (2006). Neglect and extinction: within and between sensory modalities. *Restor. Neurol. Neurosci.* 24, 217–232.

Bruno, N., Bernardis, P., and Gentilucci, M. (2008). Visually guided pointing, the Müller-Lyer illusion, and the functional interpretation of the dorsal-ventral split: conclusions from 33 independent studies. *Neurosci. Biobehav. Rev.* 32, 423–437.

Bruno, N., and Franz, V. H. (2009). When is grasping affected by the Müller-Lyer illusion? A quantitative review. *Neuropsychologia* 47, 1421–1433.

Bruno, N., Knox, P. C., and De Grave, D. D. (2010). A metanalysis of the effect of the Müller-Lyer illusion on saccadic eye movements: no general support for a dissociation of perception and oculomotor action. *Vision Res.* 50, 2671–2682.

Campbell, D. C., and Oxbury, J. M. (1976). Recovery from unilateral visuo-spatial neglect? *Cortex* 12, 303–312.

Chedru, F. (1976). Space representation in unilateral spatial neglect. *J. Neurol. Neurosurg. Psychiatr.* 39, 1057–1061.

Chokron, S., Colliot, P., Bartolomeo, P., Rhein, F., Eusop, E., Vassel, P., and Ohlmann, T. (2002). Visual, proprioceptive and tactile performance in left neglect patients. *Neuropsychologia* 40, 1965–1976.

Chokron, S., Dupierrix, E., Tabert, M., and Bartolomeo, P. (2007). Experimental remission of unilateral spatial neglect. *Neuropsychologia* 45, 3127–3148.

Costa, L. D. (1976). Interset variability on the raven coloured progressive matrices as an indicator of specific ability deficit in brain-lesioned patients. *Cortex* 12, 31–40.

Crawford, J. R., and Garthwaite, P. H. (2002). Investigation of the single case in neuropsychology: confidence limits on the abnormality of test scores and test score differences. *Neuropsychologia* 40, 1196–1208.

Cubelli, R., Nichelli, P., Bonito, V., De Tanti, A., and Inzaghi, M. G. (1991). Different patterns of dissociation in unilateral spatial neglect. *Brain Cogn.* 15, 139–159.

Daini, R., Angelelli, P., Antonucci, G., Cappa, S. F., and Vallar, G. (2002). Exploring the syndrome of spatial unilateral neglect through an illusion of length. *Exp. Brain Res.* 144, 224–237.

De Renzi, E., Faglioni, P., and Scotti, G. (1970). Hemispheric contribution to exploration of space through the visual and tactile modality. *Cortex* 6, 191–203.

Denny-Brown, D., Meyer, J. S., and Horenstein, S. (1952). The significance of perceptual rivalry resulting from parietal lesion. *Brain* 75, 433–471.

Diller, L., and Weinberg, J. (1977). "Hemi-inattention in rehabilitation. The evolution of a rational remediation program," in *Hemi-inattention and Hemisphere Specialization*, eds E. A. Weinstein and R. P. Friedland (New York: Raven Press), 62–82.

Ernst, M. O., and Bulthoff, H. H. (2004). Merging the senses into a robust percept. *Trends Cogn. Sci. (Regul. Ed.)* 8, 162–169.

Fortis, P., Maravita, A., Gallucci, M., Ronchi, R., Grassi, E., Senna, I., Olgiati, E., Perucca, L., Banco, E., Posteraro, L., Tesio, L., and Vallar, G. (2010). Rehabilitating patients with left spatial neglect by prism exposure during a visuomotor activity. *Neuropsychology* 24, 681–697.

Frassinetti, F., Bolognini, N., Bottari, D., Bonora, A., and Làdavas, E. (2005). Audiovisual integration in patients with visual deficit. *J. Cogn. Neurosci.* 17, 1442–1452.

Frassinetti, F., Bolognini, N., and Làdavas, E. (2002a). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Exp. Brain Res.* 147, 332–343.

Frassinetti, F., Pavani, F., and Làdavas, E. (2002b). Acoustical vision of neglected stimuli: interaction among spatially converging audiovisual inputs in neglect patients. *J. Cogn. Neurosci.* 14, 62–69.

Fujii, T., Fukatsu, R., Kimura, I., Saso, S., and Kogure, K. (1991). Unilateral spatial neglect in visual and tactile modalities. *Cortex* 27, 339–343.

Gainotti, G. (2010). The role of automatic orienting of attention towards ipsilesional stimuli in non-visual (tactile and auditory) neglect: a critical review. *Cortex* 46, 150–160.

Gainotti, G., Messerli, P., and Tissot, R. (1972). Qualitative analysis of unilateral spatial neglect in relation to laterality of cerebral lesions. *J. Neurol. Neurosurg. Psychiatr.* 35, 545–550.

Gallace, A., Imbornone, E., and Vallar, G. (2008). When the whole is more than the sum of the parts: evidence from visuospatial neglect. *J. Neuropsychol.* 2, 387–413.

Gangopadhyay, N., Madary, M., and Spicer, F. (2010). *Perception, Action and Consciousness. Sensorimotor Dynamics and the Two Visual Systems.* Oxford: Oxford University Press.

Gauthier, L., Dehaut, F., and Joanette, Y. (1989). The bells test: a quantitative and qualitative test for visual neglect. *Int. J. Clin. Neuropsychol.* 11, 49–54.

Gentilini, M., Barbieri, C., De Renzi, E., and Faglioni, P. (1989). Space exploration with and without the aid of vision in hemisphere-damaged patients. *Cortex* 25, 643–651.

Gilliatt, R. W., and Pratt, R. T. (1952). Disorders of perception and performance in a case of right-sided cerebral thrombosis. *J. Neurol. Neurosurg. Psychiatr.* 15, 264–271.

Gottesman, R. F., Kleinman, J. T., Davis, C., Heidler-Gary, J., Newhart, M., Kannan, V., and Hillis, A. E. (2008). Unilateral neglect is more severe and common in older patients with right hemispheric stroke. *Neurology* 71, 1439–1444.

Grigoletto, F., Zappala, G., Anderson, D. W., and Lebowitz, B. D. (1999). Norms for the Mini-Mental State Examination in a healthy population. *Neurology* 53, 315–320.

Haeske-Dewick, H. C., Canavan, A. G., and Homberg, V. (1996). Directional hyperattention in tactile neglect within grasping space. *J. Clin. Exp. Neuropsychol.* 18, 724–732.

Halligan, P. W., Fink, G. R., Marshall, J. C., and Vallar, G. (2003). Spatial cognition: evidence from visual neglect. *Trends Cogn. Sci. (Regul. Ed.)* 7, 125–133.

Heilman, K. M., Watson, R. T., and Valenstein, E. (2003). "Neglect and related disorders," in *Clinical Neuropsychology*, 4 Edn, eds K. M. Heilman and E. Valenstein (New York: Oxford University Press), 296–346.

Hier, D. B., Mondlock, J., and Caplan, L. R. (1983a). Behavioral abnormalities after right hemisphere stroke. *Neurology* 33, 337–344.

Hier, D. B., Mondlock, J., and Caplan, L. R. (1983b). Recovery of behavioral abnormalities after right hemisphere stroke. *Neurology* 33, 345–350.

Hillis, A. E., Chang, S., Heidler-Gary, J., Newhart, M., Kleinman, J. T., Davis, C., Barker, P. B., Aldrich, E., and Ken, L. (2006). Neural correlates of modality-specific spatial extinction. *J. Cogn. Neurosci.* 18, 1889–1898.

Hjaltason, H., Caneman, G., and Tegner, R. (1993). Visual and tactile rod bisection in unilateral neglect. *Cortex* 29, 583–588.

Holding, D. H. (1970). Notes and discussion. A line illusion with irrelevant depth cues. *Am. J. Psychol.* 83, 280–282.

Husain, M. (2008). "Hemispatial neglect," in *Handbook of Clinical Neurology*, eds P. J. Vinken and G. W. Bruyn (Amsterdam: Elsevier), 359–372.

Jacquin-Courtois, S., Rode, G., Pavani, F., O'shea, J., Giard, M. H., Boisson, D., and Rossetti, Y. (2010). Effect of prism adaptation on left dichotic listening deficit in neglect patients: glasses to hear better? *Brain* 133, 895–908.

Judd, C. H. (1899). A study of geometrical illusions. *Psychol. Rev.* 6, 241–261.

Kerkhoff, G. (2003). Modulation and rehabilitation of spatial neglect by sensory stimulation. *Prog. Brain Res.* 142, 257–271.

Làdavas, E., Menghini, G., and Umiltà, C. (1994). A rehabilitation study of hemispatial neglect. *Cogn. Neuropsychol.* 11, 75–95.

Leibovitch, F. S., Black, S. E., Caldwell, C. B., Ebert, P. L., Ehrlich, L. E., and Szalai, J. P. (1998). Brain-behavior correlations in hemispatial neglect using CT and SPECT: the Sunnybrook Stroke Study. *Neurology* 50, 901–908.

Mancini, F., Bolognini, N., Bricolo, E., and Vallar, G. (2011). Cross-modal processing in the occipito-temporal cortex: a TMS Study of the Müller-Lyer illusion. *J. Cogn. Neurosci.* 23, 1987–1997.

Mancini, F., Bricolo, E., and Vallar, G. (2010). Multisensory integration in the Müller-Lyer illusion: from vision to haptics. *Q. J. Exp. Psychol.* 63, 818–830.

Maravita, A., Mcneil, J., Malhotra, P., Greenwood, R., Husain, M., and Driver, J. (2003). Prism adaptation can improve contralesional tactile perception in neglect. *Neurology* 60, 1829–1831.

Mattingley, J. B., Bradshaw, J. L., and Bradshaw, J. A. (1995). The effects of unilateral visuospatial neglect on perception of Müller-Lyer illusory figures. *Perception* 24, 415–433.

Milner, A. D., and Goodale, M. (2006). *The Visual Brain in Action.* Oxford: Oxford University Press.

Mort, D. J., Malhotra, P., Mannan, S. K., Rorden, C., Pambakian, A., Kennard, C., and Husain, M. (2003). The anatomy of visual neglect. *Brain* 126, 1986–1997.

Olk, B., Harvey, M., Dow, L., and Murphy, P. J. (2001). Illusion processing in hemispatial neglect. *Neuropsychologia* 39, 611–625.

Orsini, A., Grossi, D., Capitani, E., Laiacona, M., Papagno, C., and Vallar, G. (1987). Verbal and spatial immediate memory span: normative data from 1355 adults and 1112 children. *Ital. J. Neurol. Sci.* 8, 539–548.

Perani, D., Vallar, G., Cappa, S., Messa, C., and Fazio, F. (1987). Aphasia and neglect after subcortical stroke. A clinical/cerebral perfusion correlation study. *Brain* 110, 1211–1229.

Pizzamiglio, L., Antonucci, G., Judica, A., Montenero, P., Razzano, C., and Zoccolotti, P. (1992). Cognitive rehabilitation of the hemineglect disorder in chronic patients with unilateral right brain damage. *J. Clin. Exp. Neuropsychol.* 14, 901–923.

Pizzamiglio, L., Guariglia, C., Antonucci, G., and Zoccolotti, P. (2006). Development of a rehabilitative program for unilateral neglect. *Restor. Neurol. Neurosci.* 24, 337–345.

Ro, T., and Rafal, R. D. (1996). Perception of geometric illusions in hemispatial neglect. *Neuropsychologia* 34, 973–978.

Rode, G., Klos, T., Courtois-Jacquin, S., Rossetti, Y., and Pisella, L. (2006a). Neglect and prism adaptation: a new therapeutic tool for spatial cognition disorders. *Restor. Neurol. Neurosci.* 24, 347–356.

Rode, G., Michel, C., Rossetti, Y., Boisson, D., and Vallar, G. (2006b). Left size distortion (hyperschematia) after right brain damage. *Neurology* 67, 1801–1808.

Ronchi, R., Posteraro, L., Fortis, P., Bricolo, E., and Vallar, G. (2009). Perseveration in left spatial neglect: drawing and cancellation tasks. *Cortex* 45, 300–312.

Rorden, C., and Brett, M. (2000). Stereotaxic display of brain lesions. *Behav. Neurol.* 12, 191–200.

Rorden, C., Fruhmann Berger, M., and Karnath, H. O. (2006). Disturbed line bisection is associated with posterior brain lesions. *Brain Res.* 1080, 17–25.

Schenkenberg, T., Bradford, D. C., and Ajax, E. T. (1980). Line bisection and unilateral visual neglect in patients with neurologic impairment. *Neurology* 30, 509–517.

Schindler, I., Clavagnier, S., Karnath, H. O., Derex, L., and Perenin, M. T. (2006). A common basis for visual and tactile exploration deficits in spatial neglect? *Neuropsychologia* 44, 1444–1451.

Schroder, A., Wist, E. R., and Hömberg, V. (2008). TENS and optokinetic stimulation in neglect therapy after cerebrovascular accident: a randomized controlled study. *Eur. J. Neurol.* 15, 922–927.

Smania, N., and Aglioti, S. (1995). Sensory and spatial components of somaesthetic deficits following right brain damage. *Neurology* 45, 1725–1730.

Stein, B. E., and Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nat. Rev. Neurosci.* 9, 255–266.

Vallar, G. (1998). Spatial hemineglect in humans. *Trends Cogn. Sci. (Regul. Ed.)* 2, 87–97.

Vallar, G. (2000). "The methodological foundations of human neuropsychology: studies in brain-damaged patients," in *Handbook of Neuropsychology*, 2nd Edn, eds F. Boller, J. Grafman, and G. Rizzolatti (Amsterdam: Elsevier), 305–344.

Vallar, G., Antonucci, G., Guariglia, C., and Pizzamiglio, L. (1993a). Deficits of position sense, unilateral neglect and optokinetic stimulation. *Neuropsychologia* 31, 1191–1200.

Vallar, G., Bottini, G., Rusconi, M. L., and Sterzi, R. (1993b). Exploring somatosensory hemineglect by vestibular stimulation. *Brain* 116, 71–86.

Vallar, G., Bottini, G., Sterzi, R., Passerini, D., and Rusconi, M. L. (1991a). Hemianesthesia, sensory neglect, and defective access to conscious experience. *Neurology* 41, 650–652.

Vallar, G., Rusconi, M. L., Geminiani, G., Berti, A., and Cappa, S. F. (1991b). Visual and nonvisual neglect after unilateral brain lesions: modulation by visual input. *Int. J. Neurosci.* 61, 229–239.

Vallar, G., and Daini, R. (2006). "Visual perceptual processing in unilateral spatial neglect: the case of visual illusions," in *Imagery and Spatial Cognition: Methods, Models and Cognitive Assessment*, eds T. Vecchi and G. Bottini (Amsterdam/Philadelphia: John Benjamins Publishing Company), 337–362.

Vallar, G., Guariglia, C., Magnotti, L., and Pizzamiglio, L. (1995a). Optokinetic stimulation affects both vertical and horizontal deficits of position sense in unilateral neglect. *Cortex* 31, 669–683.

Vallar, G., Guariglia, C., Nico, D., and Bisiach, E. (1995b). Spatial hemineglect in back space. *Brain* 118, 467–472.

Vallar, G., Guariglia, C., and Rusconi, M. L. (1997). "Modulation of the neglect syndrome by sensory stimulation," in *Parietal Lobe Contributions to Orientation in 3D Space*, eds P. Thier and H.-O. Karnath (Heidelberg: Springer-Verlag), 555–578.

Vallar, G., and Mancini, F. (2010). "Mapping the neglect syndrome onto neurofunctional streams," in *Perception, Action, and Consciousness. Sensorimotor Dynamics and Two Visual Systems*, eds N. Gangopadhyay, M. Madary, and F. Spicer (Oxford: Oxford University Press), 183–215.

Vallar, G., Rusconi, M. L., Bignamini, L., Geminiani, G., and Perani, D. (1994a). Anatomical correlates of visual and tactile extinction in humans: a clinical CT scan study. *J. Neurol. Neurosurg. Psychiatry* 57, 464–470.

Vallar, G., Rusconi, M. L., Fontana, S., and Musicco, M. (1994b). Tre test di esplorazione visuo-spaziale: taratura su 212 soggetti normali. *Arch. Psicol. Neurol. Psichiatr.* 55, 827–841.

Vallar, G., Sterzi, R., Bottini, G., Cappa, S., and Rusconi, M. L. (1990). Temporary remission of left hemianesthesia after vestibular stimulation. A sensory neglect phenomenon. *Cortex* 26, 123–131.

Verdon, V., Schwartz, S., Lovblad, K. O., Hauert, C. A., and Vuilleumier, P. (2010). Neuroanatomy of hemispatial neglect and its functional components: a study using voxel-based lesion-symptom mapping. *Brain* 133, 880–894.

Villardita, C. (1987). Tactile exploration of space and visual neglect in brain-damaged patients. *J. Neurol.* 234, 292–297.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 28 July 2011; accepted: 01 November 2011; published online: 22 November 2011.*

*Citation: Mancini F, Bricolo E, Mattioli FC and Vallar G (2011) Visuo-haptic interactions in unilateral spatial neglect: the cross modal Judd illusion. Front. Psychology* ***2****:341. doi: 10.3389/fpsyg.2011.00341*

*This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.*

frontiers in
PSYCHOLOGY

REVIEW ARTICLE
published: 06 May 2011
doi: 10.3389/fpsyg.2011.00089

# Studying multisensory processing and its role in the representation of space through pathological and physiological crossmodal extinction

***Stéphane Jacobs*[1,2,3]*, *Claudio Brozzoli*[4], *Fadila Hadj-Bouziane*[1,2,3], *Martine Meunier*[1,2,3] and *Alessandro Farnè*[1,2,3]***

[1] *ImpAct Team, INSERM U1028, Lyon Neuroscience Research Center, Lyon, France*
[2] *ImpAct Team, CNRS UMR5292, Lyon Neuroscience Research Center, Lyon, France*
[3] *University Lyon I, Lyon, France*
[4] *Brain, Body and Self Laboratory, Department of Neuroscience, Karolinska Institutet, Stockholm, Sweden*

***Edited by:***
*Nadia Bolognini, University of Milano-Bicocca, Italy*
*Angelo Maravita, University of Milano-Bicocca, Italy*

***Reviewed by:***
*Francesca Frassinetti, University of Bologna, Italy*
*Elena Natale, University of Milano-Bicocca, Italy*

****Correspondence:***
*Stéphane Jacobs and Alessandro Farnè, INSERM U1028, CNRS UMR5292, Centre de Recherche en Neurosciences de Lyon, Equipe IMPACT, 16 Avenue du Doyen Lépine, 69500 Bron, France.*
*email: stephane.jacobs@inserm.fr; alessandro.farne@inserm.fr*

The study of crossmodal extinction has brought a considerable contribution to our understanding of how the integration of stimuli perceived in multiple sensory modalities is used by the nervous system to build coherent representations of the space that directly surrounds us. Indeed, by revealing interferences between stimuli in a disturbed system, extinction provides an invaluable opportunity to investigate the interactions that normally exist between those stimuli in an intact system. Here, we first review studies on pathological crossmodal extinction, from the original demonstration of its existence, to its role in the exploration of the multisensory neural representation of space and the current theoretical accounts proposed to explain the mechanisms involved in extinction and multisensory competition. Then, in the second part of this paper, we report recent findings showing that physiological multisensory competition phenomena resembling clinical crossmodal extinction exist in the healthy brain. We propose that the development of a physiological model of sensory competition is fundamental to deepen our understanding of the cerebral mechanisms of multisensory perception and integration. In addition, a similar approach to develop a model of physiological sensory competition in non-human primates should allow combining functional neuroimaging with more invasive techniques, such as transient focal lesions, in order to bridge the gap between works done in the two species and at different levels of analysis.

**Keywords: extinction, multisensory, neuropsychology, perception, peripersonal space, non-human primate**

## INTRODUCTION

As in many fields of neuropsychology, the study of multisensory processing dysfunctions in brain-lesioned patients has been of invaluable help in better understanding the mechanisms and functional roles of multisensory perception and integration in human cognition. In particular, crossmodal extinction has constituted a very useful model for investigating how multiple representations of the space that surrounds us are built in our brains through the integration of sensory information perceived in different modalities (Làdavas and Farnè, 2004; Farnè et al., 2007).

Extinction is a neuropsychological syndrome closely related to spatial neglect[1], both arising in a vast majority of cases following damage to the right hemisphere, most typically in the posterior parietal region (Becker and Karnath, 2007). While neglect patients fail to consciously perceive sensory events located in the contralesional side of space (i.e., most commonly the left hemispace; Driver and Vuilleumier, 2001; Kerkhoff, 2001) extinction patients have difficulties in consciously perceiving a contralesional stimulus only when it is delivered simultaneously with an ipsilesional stimulation (Bender, 1952). This deficit specific to conditions of double simultaneous stimulation (DSS), together with a (at least almost) normal detection of contralesional stimuli presented in isolation, is the hallmark of extinction. Besides the widely recognized spatial nature of these two syndromes, extinction and neglect patients also exhibit deficits in the temporal processing of contralesional stimuli (Husain et al., 1997; Rorden et al., 1997). In particular, extinction patients exhibit an abnormally long attentional dwell-time (i.e., the period during which the perception of a first attended stimulus interferes with that of a subsequent one) in the contralesional visual hemifield, and this might contribute to the competitive bias in favor of ipsilesional stimuli in these patients (di Pellegrino et al., 1998). However, the temporal order of ipsi- and contralesional stimuli seems not to influence extinction rate, which remains maximal for simultaneous stimuli (di Pellegrino et al., 1997a; Rorden et al., 2009).

Neglect and extinction can affect all sensory modalities, either separately, across patients, or jointly, in the same individual (Brozzoli et al., 2006). Extinction was first reported in humans in the tactile

---

[1] Although this is beyond the scope of this review, it should be noted that the question of whether extinction and neglect should be conceived of as common or separate deficits is still debated. Indeed, while extinction is often considered as a milder form or a residual manifestation of neglect after recovery (e.g., Heilman et al., 1993), dissociations have been reported (e.g., di Pellegrino and De Renzi, 1995; Cocchini et al., 1999; Geeraerts et al., 2005) and distinct neural substrates have been proposed (Vallar et al., 1994; Karnath et al., 2003), challenging the view that both syndromes are just different levels of impairment along the same continuum.

modality by Oppenheim (1885; cited in Benton, 1956). The phenomenon has subsequently been described for visual (Poppelreuter, 1917; cited in Bender and Teuber, 1946; Critchley, 1953; Heilman et al., 1993) and auditory (Bender, 1952; Heilman et al., 1970; De Renzi et al., 1984; Jacquin-Courtois et al., 2010) stimuli as well. Cases of olfactory (Mesulam, 1981; Bellas et al., 1988a,b) and gustatory (Bender and Feldman, 1952) extinction have also been reported, although it is unclear whether they reflected true olfactory or gustatory extinction, or were rather due to somatosensory deficits (Kobal et al., 1989; Berlucchi et al., 2004).

The first observation of extinction in animals also dates back to 1885. In parallel to Oppenheim's discovery in humans, Loeb reported that dogs with damage to one hemisphere favored pieces of meat situated on the side of the lesion only upon simultaneous presentation of two pieces of meat (Benton, 1956). Since then, a few studies have sought to develop a model of neglect or extinction in non-human primates, using lesions of parietal, frontal, or superior temporal cortices to induce behaviors similar to those observed in patients. Authors have investigated the impact of these lesions on the perception of visual (Faugier-Grimaud et al., 1978; Lynch and McLaren, 1989; Watson et al., 1994; Heilman et al., 1995) or tactile (Ettlinger and Kalsbeck, 1962) stimuli, or both (but as separate modalities; Matelli et al., 1982; Rizzolatti et al., 1983). For example, Lynch and McLaren (1989) demonstrated visual extinction in five monkeys with unilateral parietal damage trained to detect single or double small visual targets presented at various eccentricities. Although still able to direct saccades toward a single contralesional stimulus, all five monkeys were unable to detect the same target when presented simultaneously with an ipsilesional one.

For about a century after its discovery, extinction was thought to occur only within sensory modalities, despite anecdotal reports of crossmodal extinction (Bender, 1952). As a unimodal phenomenon, it provided insights into the neural bases of unisensory competition (Kinsbourne, 1977; Duncan, 1996; Riddoch et al., 2009). However, when its multisensory dimension became clear, extinction could then be used as a model to study the multisensory neural representation of space. The present review focuses on this most recent development in extinction research and the considerable insight it brought into human spatial cognition, before proposing new developments in healthy individuals and non-human primates as the support for future research on extinction. The first sections provide a brief outline of the original demonstration of the existence of multisensory (or crossmodal) extinction, its role in the investigation of multisensory spatial representation, and an overview of the current theoretical accounts of extinction. In the second part of this paper, we propose two future developments for extinction research that we believe are key to deepen our understanding of the cerebral mechanisms of extinction and multisensory processing. The first one is the exploration of physiological extinction phenomena in healthy humans, their behavioral mechanisms and their neural underpinnings, thanks to functional neuroimaging and brain stimulation methods. The second one is the refinement of the existing non-human primate model of extinction by using the same behavioral approach as in humans to strengthen the validity for inter-species comparisons, also building on the unique opportunity provided by animal research of combining focal lesions techniques to functional neuroimaging methods.

## EXTINCTION AS A MULTISENSORY PHENOMENON

Posner's seminal work on selective spatial attention in the 1980s (Posner et al., 1982, 1984, 1987a,b) motivated the first attempts to systematically investigate whether stimuli presented in different sensory modalities could compete for common, supramodal attentional resources. The first evidence of multisensory competition for attentional resources was brought by Farah et al. (1989), who used a modified version of Posner's task (Posner et al., 1982). In this study, a lateralized auditory cue, either valid or invalid, preceded the presentation of a visual target either to the left or to the right of a central fixation point, and participants had to respond as soon as possible to the visual target by pressing the corresponding button. Patients with right parietal lesions were slower at detecting contralesional left visual targets when previously invalidly cued to the right side of space, suggesting a deficit in disengaging spatial attention from the wrongly cued ipsilesional location to subsequently relocate it to the contralesional visual target. Farah et al. (1989) interpreted these results as evidence for the existence of a supramodal system controlling the allocation of spatial attention. However, auditory cues did not interfere with the perception of visual stimuli to the point of hampering their detection. While these lengthened reaction times indeed reflected interferences between the processing of auditory cues and visual target stimuli, they remained much less spectacular than the outright omissions of stimulations observed in extinction patients during DSS.

Along the same line, Inhoff et al. (1992) investigated more explicitly the possible existence of crossmodal, visuotactile extinction in patients suffering from right parieto-frontal lesions and exhibiting unimodal visual and tactile extinction. To this end, they used the classical confrontation test, which consists in having the patient face the experimenter and apply stimuli either to the left or right side of space, or both simultaneously. Typically, the patient is asked to verbally report which side(s) the stimulation was applied to, and the experimenter compares performance on single stimulus detection with responses to DSS. As detailed previously, the hallmark of extinction is a (relatively) normal detection of single stimulations together with a deficit in detecting contralesional stimuli on DSS trials. In Inhoff et al.'s (1992) study, visual stimuli corresponded to wiggling either index finger or both at the patient's eye level, while tactile stimulations were applied by lightly touching the dorsum of the hand(s) while the patient had the eyes closed. The three patients tested showed pronounced extinction within the visual and tactile modalities, but no sign of visuotactile extinction, which led Inhoff et al. (1992) to conclude that processing of sensory inputs from both modalities is controlled by separate spatial attention systems, and that parieto-frontal brain structures may not contribute to the elaboration of a supramodal representation of space.

These findings were subsequently challenged by abundant neurophysiological data collected in the monkey that demonstrated the existence of multimodal neurons in areas commonly associated with spatial attention and extinction, such as parietal cortex (Duhamel et al., 1998) and basal ganglia (Graziano and Gross, 1993), as well as ventral premotor cortex (Rizzolatti et al., 1981; Graziano et al., 1994). These neurons respond to stimuli delivered in different sensory modalities (e.g., vision and touch) and have been suggested to constitute the substrate for crossmodal integration and multisensory representation of space. The most widely studied cells are visuotactile neurons, which have a visual receptive field in register with their tactile receptive field, that is, they respond to visual stimuli applied on or in the space immediately surrounding a specific body part, thus providing multiple multisensory maps of space centered on different body parts. Importantly, this implies that visuotactile interactions are strongest near the body, but fade as the visual stimulus moves away.

In light of these findings, Inhoff et al.'s (1992) failure to demonstrate the existence of visuotactile extinction might just be due to the fact that their visual stimulations were delivered too far from the hand, at patients' eye level, where they would interact much less, if at all, with the tactile stimuli. Consistent with this interpretation, di Pellegrino et al. (1997b) and Mattingley et al. (1997), independently and at about the same time, tested patients with right brain damage and suffering from unisensory extinction in the tactile or visual modality, or both. Both investigations were based on a confrontation test similar to that used in Inhoff et al.'s (1992) investigation, except that visual stimuli were applied near the hand. In both studies, patients exhibited strong visuotactile extinction, with a visual stimulus applied near the ipsilesional hand hampering the detection of a tactile stimulation of the contralesional hand. The reverse pattern of extinction (i.e., a right tactile stimulus extinguishing a left visual stimulus applied near the hand) was also present when investigated (Mattingley et al., 1997). Importantly, the patient tested in di Pellegrino et al.'s (1997b) study did not show visuotactile extinction when the ipsilesional visual stimulus was applied far from the hand, supporting the hypothesis that these multisensory interactions are supported by mechanisms similar to those described in the monkey (Rizzolatti et al., 1981; Graziano et al., 1997; Duhamel et al., 1998). However, the fact that crossmodal extinction was still present when the visual stimulus was applied at patients' eye level in Mattingley et al.'s (1997) study suggests that these interactions, although strongest in near space, are not an all-or-nothing phenomenon. This is consistent with subsequent observations that, although usually milder, crossmodal extinction can still occur with ipsilesional stimuli delivered far from the hand (Làdavas et al., 1998a; Farnè and Làdavas, 2000), suggesting that the transition from near to far space is gradual rather than an abrupt boundary at arm's length (Longo and Lourenco, 2006).

Since the demonstration of the existence of visuotactile extinction, many investigations have used extinction as a window on the mechanisms of normal multisensory processing in the human brain, showing for example that other sensory modalities can interact, such as touch and audition (Làdavas et al., 2001; for a similar demonstration using alloesthesia, see Ortigue et al., 2005). Most importantly, as we will detail in the next section, the study of neglect and crossmodal extinction has proven an invaluable tool for the exploration of human spatial cognition.

## MULTISENSORY EXTINCTION AND THE MODULAR REPRESENTATION OF SPACE

Although our conscious experience of the space that surrounds us is that of a unitary entity, it is actually the product of the integration of multiple neural representations, subserved by distinct cerebral systems and involving the convergence of information perceived in different sensory modalities (Calvert et al., 2004; Spence and Driver, 2004). The first evidence of separate representations of distinct sectors of space in humans has been brought by the demonstration that neglect can be restricted to personal space (e.g., Bisiach et al., 1986) or to the space within (Halligan and Marshall, 1991; Berti and Frassinetti, 2000) or beyond (Vuilleumier et al., 1998) reaching distance. This distinction between the representations of near and far space had been previously shown in monkeys following focal lesions to the postarcuate cortex (area 6), which induced tactile and visual neglect for the contralateral space immediately surrounding the body, and to the frontal eye field (area 8), which gave rise to visual neglect for far space, with a tendency to deviate the eyes and the head toward the side of the lesion and to decrease spontaneous eye movements contralateral to the lesion, as typically observed in human neglect (Rizzolatti et al., 1983). Subsequently however, the study of patients with crossmodal extinction has provided a much more detailed picture of the modularity of the cerebral representation of space.

Indeed, the first report of spatially selective visuotactile extinction described above (di Pellegrino et al., 1997b) provided evidence that the space located within reaching distance, also often termed peripersonal space (for a detailed discussion of these concepts of spatial representation, see Cardinali et al., 2009), can be divided into two parts: a near sector, constituted of the space immediately surrounding body (about 5 cm), where visuotactile integration is strongest; and a far sector (at approximately 35 cm from the body), where visuotactile interactions are weaker. In addition, these studies showed that the maps of near peripersonal space were centered on a specific body part, that is, in those cases, the hand. In particular, in one of their control conditions, di Pellegrino et al. (1997b) presented the visual stimulus at the same spatial location as in the regular confrontation test (i.e., next to where the patient's hand used to be), while the patient held his hand behind his back. In this condition, the patient detected flawlessly the left tactile stimulus when it was delivered concurrently with the ipsilesional visual stimulus, showing that the visuotactile interaction evidenced in the normal confrontation test occurred in hand-centered space, and not in a retinal or other more general egocentric (e.g., head- or trunk-centered) spatial frame of reference. This interpretation was further supported by the observation that, while a visual stimulus might prevent the detection of a contralesional tactile stimulus if it is applied near the ipsilesional hand in a patient with tactile extinction, it can on the contrary enhance the detection of the same contralesional tactile stimulus if it is instead applied near the contralesional hand during tactile DSS (Làdavas et al., 1998a). Again, this modulation of tactile perception by visual stimuli was dramatically diminished if the visual stimulus was applied far from the hand. Therefore, in the former case, the simultaneous activation of the visuotactile representations of the left and right hands by the tactile and visual stimulus, respectively, leads to the extinction of the weaker representation supported by the lesioned right hemisphere, that is, the one of the left hand. By contrast, in the latter case, when the visual stimulus is applied near the left hand, it enhances the activation of the weaker representation, thereby allowing the detection of a left tactile stimulus even when a concurrent tactile stimulus is delivered on the right hand.

Subsequently, similar paradigms have been used to reveal the existence of multisensory maps of near peripersonal space and their relation to other body parts. In particular, crossmodal extinction has been used to demonstrate interactions between vision and touch in a space centered on the face in humans (Làdavas et al., 1998b). In this study, extinction patients received unilateral and bilateral tactile stimuli on the cheeks, together with a visual stimulus applied either on the ipsilesional or on the contralesional side. Akin to what has been described for the hand, the visual stimulus extinguished the contralesional tactile stimulus when presented ipsilesionally, whereas it enhanced its detection when delivered near the contralesional cheek. Again, visual stimuli presented far from the face modulated touch perception to a much lesser extent. Further investigation strengthened the evidence of a modular organization of visuotactile near peripersonal space, by directly studying the effects of ipsilesional visual stimuli delivered to homologous or non-homologous body parts (i.e., the face or the hand) to that receiving the contralesional tactile stimulus (Farnè et al., 2005a). Visuotactile extinction observed in near peripersonal space was stronger when homologous body parts were stimulated, while it was weak and comparable in the homologous and non-homologous condition in far peripersonal space. Near peripersonal space is thus not represented as a unitary entity, but rather consists of different modules, separately representing multisensory space, possibly centered on several body parts.

Finally, similar modulations of tactile extinction have been revealed through multisensory interactions between touch and audition (Làdavas et al., 2001; Farnè and Làdavas, 2002). Detection of contralesional single touches applied on the neck of tactile extinction patients was prevented by ipsilesional white noise stimuli delivered near the patients' head. When the auditory stimulus was delivered far from the head, audiotactile extinction was dramatically reduced. Interestingly, spatially selective interactions between audition and touch were strongest when the auditory stimuli came from the back, rather than from the front, of the patients' head. Additionally, the spatial specificity of the modulation of touch perception by audition was more pronounced for white noise bursts than pure tones. This suggests that different degrees of multisensory integration may occur depending on the complexity of the stimuli, and on the functional relevance of a given sensory modality for a particular sector of space (Farnè and Làdavas, 2002).

Together, these phenomena of crossmodal extinction and facilitation show that multisensory maps of near peripersonal space are elaborated through the integration of multisensory information, presumably at the neuronal level, via multimodal cells similar to those described in the monkey (Rizzolatti et al., 1981; Graziano and Gross, 1993; Graziano et al., 1994, 1997; Duhamel et al., 1998). As we have already mentioned, those neurons have been found in several cortical regions in the monkey that are commonly associated to spatial attention and extinction (Vallar et al., 1994; Karnath et al., 2003; Corbetta et al., 2005), including parietal (Duhamel et al., 1998) and ventral premotor (Rizzolatti et al., 1981; Graziano et al., 1994) cortices. This interpretation is supported by the fact that crossmodal extinction in humans and the activity of multimodal neurons in monkeys are modulated by similar parameters. Indeed, multimodal neurons, whether visuotactile (Rizzolatti et al., 1981; Graziano and Gross, 1993; Graziano et al., 1994; Duhamel et al., 1998) or audiotactile (Graziano et al., 1999), progressively cease to fire when the non-corporeal stimulus (i.e., visual or auditory) moves away from the relevant body part. Furthermore, similarly to audiotactile extinction described in patients (Farnè and Làdavas, 2002), the activity of monkey audiotactile neurons is also modulated by auditory complexity (Graziano et al., 1999). Finally, tool use induces plastic modifications of the multisensory representation of near space in both monkeys and humans. In monkeys, visuotactile neurons that normally fire only when visual stimuli are applied near the hand show an enlarged visual receptive field after tool use, which encompasses the hand and the tool (Iriki et al., 1996). In patients, visuotactile extinction increases strongly in the region of space that was previously beyond reach (i.e., far space) but made accessible by the use of the tool (for a similar effect in neglect patients, see Berti and Frassinetti, 2000; in healthy individuals, see Maravita et al., 2002; Holmes et al., 2004; for review, see Maravita and Iriki, 2004; Farnè et al., 2007). Additionally, this remapping occurs only after active and functional use of the tool in both species (Iriki et al., 1996; Maravita et al., 2002; Farnè et al., 2005b).

In sum, the body of evidence reviewed here and collected through the study of crossmodal extinction supports the hypothesis of a cerebral representation of multisensory space organized in a modular fashion, supported by several neural systems devoted to the representation of different sectors of space, in different coordinates, and probably for different behavioral purposes (Stein and Arigbede, 1972).

## THEORETICAL ACCOUNTS OF UNI- AND MULTISENSORY EXTINCTION

As stated earlier, extinction and neglect are often considered manifestations of the same underlying deficit, representing different levels of impairment along a continuum (e.g., Heilman et al., 1993). As such, both disorders have usually been addressed jointly by several theories proposed to account for their behavioral manifestations. Whether they postulate a deficit in the representation of the contralesional hemispace (Bisiach et al., 1981) or in the ability to disengage attention from ipsilesional stimuli (Posner et al., 1984), a deviation (Jeannerod and Biguer, 1987) or a distortion (Bisiach et al., 1996) of spatial representation, most of these interpretations have been based on a dominant role of the right hemisphere in representing and orienting attention to both sides of space, while the left hemisphere would only be concerned with the right contralateral hemispace (Heilman et al., 1993, 1997). While this assumption might explain the spatial bias exhibited by neglect patients who fail to attend to or act toward stimuli in the contralesional hemispace, it is hard to reconcile with the competitive nature of extinction. Indeed, the hypothesis that the right lesion would leave intact only the attentional capacities of the left hemisphere leads to the prediction of an inability to detect single contralesional stimuli, rather than extinction arising from the competition between ipsi- and contralesional stimuli for attentional resources (for a detailed discussion of this point, see di Pellegrino and De Renzi, 1995). In fact, this view is hardly compatible with the competitive aspects of neglect itself, which has been shown to be reduced for example on a cancelation task when ipsilesional targets were progressively removed (Mark et al., 1988). This finding has been interpreted as evidence for a deficit in disengaging attention from ipsilesional stimuli, but might also be conceived of as suggesting competitive underlying mechanisms.

An alternative model addresses explicitly the competitive dimension of extinction (and neglect) by proposing that these phenomena result from a breakdown in the dynamic balance that normally exists in the reciprocal inhibition between homologous areas of the two hemispheres that orient spatial attention in opposing, contralateral directions (Kinsbourne, 1977, 1987). According to this proposal, a lesion lateralized to one hemisphere induces not only its hypoactivity, but also hyperactivity in the other hemisphere due to the release of inhibition by the affected one. This imbalance generates a general bias toward the ipsilesional side of space, which results in a biased competition between simultaneous stimuli (Duncan, 1996; Driver et al., 1997). The higher frequency of neglect and extinction following left than right brain damage would be explained by the fact that the rightward orienting bias generated by the left hemisphere is stronger than the opposing bias induced by the right hemisphere. This model provides a convincing explanation of why, in extinction, a contralesional stimulus is normally detected most of the time when delivered in isolation, but goes undetected when presented simultaneously with an ipsilesional stimulus (Bender, 1952). Because the deficit lies at the level of interhemispheric rivalry, in the former case, the lesioned hemisphere is still capable of processing (almost normally) the isolated contralesional stimulus. In the latter case, however, the concurrent activation of the intact hemisphere by the ipsilesional stimulus leads to enhanced inhibition of homologous areas on the side of the lesion, thereby preventing the conscious detection of the contralesional stimulus. Strong support to this model comes from brain stimulation studies using repetitive transcranial magnetic stimulation (rTMS) to temporarily decrease the activity of the *intact* hemisphere in brain-damaged patients suffering from tactile extinction (Oliveri et al., 1999) and neglect (Oliveri et al., 2001). In both cases, transient deactivation of the healthy hemisphere temporarily decreased the manifestations of extinction or neglect, in accordance with the predictions drawn from Kinsbourne's theory. Although the implications of this model have usually been discussed in the framework of unimodal extinction (e.g., di Pellegrino and De Renzi, 1995), it is applicable to the case of crossmodal extinction as well, whether the competition is considered to happen at the level of the orientation of some supramodal attentional system, or at a more physiological level where multisensory representations of near peripersonal space interact directly through reciprocal transcallosal connections.

The results of recent neuroimaging studies have led to a more complete picture of the pathophysiological mechanisms involved in neglect and extinction, and given further support to their competitive accounts. Corbetta et al. (2000) orienting of spatial visual attention toward a specific location involves a bilateral, dorsal frontoparietal network comprising the intra-parietal sulcus and superior parietal lobule, and the frontal eye field. Another, more ventral network, constituted of the temporoparietal and inferior parietal region and the caudal part of the inferior frontal gyrus, redirects attention toward novel stimuli by sending signals to the dorsal network. This ventral network is lateralized to the right hemisphere, and corresponds to the location of lesions usually associated with neglect. Further work in right brain-lesioned patients with neglect has shown that lesions in these patients not only directly affect the ventral attention network, but also indirectly modifies the activity within the structurally intact bilateral dorsal network (Corbetta et al., 2005). Indeed, the disruption of the normal interaction between the ventral and dorsal systems resulted in these patients in abnormally unbalanced activity in the left and right dorsal parietal cortices, with increased activations in the contralesional left hemisphere and comparatively reduced activity on the right side. Further, the hyperactivity of the left superior parietal lobe correlated with the patients' bias toward the ipsilesional side of space. These findings thus provide strong support to the competitive accounts of neglect and extinction, and suggest that dorsal parietal cortex is the site of the interhemispheric activity imbalance responsible for neglect and extinction patients' rightward spatial bias.

In the framework offered by these models, it is possible to make predictions about healthy individuals' behavior. Indeed, the interhemispheric rivalry hypothesis suggests that even in a healthy brain, competition occurs between simultaneous stimuli. In line with this hypothesis, previous work has shown, using positron emission tomography (PET), that simultaneous bilateral visual stimulation induces a reduction in response of the primary visual cortex when compared to the activation observed after contralateral single stimulation, thereby providing physiological support for the existence of interhemispheric rivalry occurring in the intact brain at early levels of processing (Fink et al., 2000). Thus, it should be possible to elicit extinction in healthy subjects performing a detection task with simple-feature stimuli, provided that the system is placed in challenging enough conditions. However, in contrast with patients' behavior, healthy individuals should exhibit no spatial bias in favor of one side or the other. Indeed, while Kinsbourne postulated that the higher occurrence of neglect and extinction following right brain damage was due to a stronger rightward attentional orienting bias generated by the left hemisphere, he suggested that, in the absence of lesion, "the right brain's opposing leftward bias […] holds the left hemisphere's rightward bias in check" (Kinsbourne, 1987, p. 72). This predicts that no spatial bias should be observed in the physiological state. Further, according to Corbetta et al.'s (2005) model, imbalance between the activity of left and right dorsal parietal cortices only occurs following lesion to the ventral attention network, lateralized to the right hemisphere.

Therefore, the competition should be balanced between the two hemispheres in subjects in a normal state of vigilance (Matthias et al., 2009; see also Çiçek et al., 2007).

## CROSSMODAL EXTINCTION AS A PHYSIOLOGICAL PHENOMENON

The idea that extinction observed in brain-damaged patients may be a pathological exacerbation of limits to the physiological capacities of the brain is not new. Bender himself, in his seminal monograph about disorders of perception (Bender, 1952), anecdotally reported what resembles extinction-like phenomena in normal subjects during tactile DSS on non-homologous body parts. Subjects frequently missed the first of a series of DSS when they were not aware of what body parts would be stimulated. This observation led Bender to assume that more consistent patterns of extinction-like behavior might be observed in healthy individuals with the appropriate testing paradigm and devices. More recent works have investigated competitive phenomena in healthy subjects in the visual (Duncan, 1984; Gorea and Sagi, 2002) and tactile (Meador et al., 2001; Marcel et al., 2004; Serino et al., 2008) domain, but the induction of extinction-like patterns of performance required the use of complex experimental displays and demands (discrimination task, unbalanced stimuli, masking procedures, etc…) that are highly likely to rely on different processes than the mere detection of simultaneous simple-feature stimuli. Other groups were able to elicit extinction-like behaviors in healthy individuals, but induced to this purpose temporary and reversible "lesions" using either TMS or transcranial direct current stimulation (tDCS) to transiently impair sensory processing (Dambeck et al., 2006; Meister et al., 2006; Sparing et al., 2009; for review, see Oliveri and Caltagirone, 2006). Below, we will review recent data from our group and others showing that extinction-like phenomena can be elicited in healthy individuals without brain stimulation, using simple stimuli and tasks much more similar to the classical confrontation test, provided the appropriate experimental conditions are met, as Bender had predicted.

In accordance with the predictions drawn from the competitive accounts of extinction, our group sought to demonstrate the existence of tactile extinction in healthy individuals (Farnè et al., 2007). The hypothesis was that the somatosensory system is intrinsically limited when simultaneous stimuli compete to reach consciousness, so that extinction should manifest itself physiologically, provided the stimuli are challenging enough to perceive. To address this last point, the intensity of stimuli was first titrated independently on each side in order for the subject to detect 90% of left and right single stimulations, in order to avoid ceiling effects. During the experiment, neurologically healthy subjects performed a task as close as possible to the confrontation test, to investigate the same processes as those involved in pathological extinction. Specifically, participants were asked to localize tactile stimuli delivered to either or both index fingers by verbally responding "left," "right," "both," or "none" as fast as they could, while their hands were out of their sight. While subjects were able to report approximately 90% of single stimulations, as expected from the titration procedure, the rate of correct detection dropped to 78% for DSS. This score was lower than the joint probability of detecting both stimuli, confirming that this drop in performance was due to a genuine extinction phenomenon. Finally, in agreement with the prediction made by the hemispheric rivalry hypothesis, there was no spatial bias in the distribution of participants' errors, meaning that they missed equally often left and right stimuli when a stimulus was extinguished on DSS trials.

This study is the first demonstration of the existence of tactile extinction-like phenomena in healthy individuals, using a very simple task, as close as possible to the classical confrontation test used with patients, as well as simple-feature stimuli. This allows us to confidently believe that we are studying the very same processes that are at play in clinical extinction. We are currently investigating whether the same procedure applied to visual stimuli results in patterns of performance reminiscent of extinction patients' behavior. Preliminary data suggest that the observations made for touch hold for the visual modality. Indeed, participants miss stimuli on DSS trials while they perceive single stimulations almost perfectly and, again, do not seem to exhibit a spatial bias when failing to perceive one of two simultaneous stimuli.

Multisensory interactions in the intact brain have already been studied in many previous works. Of particular relevance to the question addressed here are studies about the Colavita effect, a phenomenon reflecting the fact that vision dominates over other sensory modalities in most experimental conditions. Colavita (1974) originally designed an experiment in which visual, auditory, or audiovisual stimuli were randomly presented, asking participants to report what they had perceived. Typically, on a certain number of bimodal trials, subjects missed the auditory stimulus and reported only the visual one, while they had no difficulties in reporting isolated stimuli. There has been lately a resurgence of interest in this phenomenon, and many studies have confirmed the existence of this effect under more controlled experimental conditions (for review, see Spence, 2009). While this effect is somewhat reminiscent of the manifestations of crossmodal extinction, it lacks a spatial dimension for the parallel with extinction and interhemispheric competition to be solid. Recently however, this effect of visual dominance has been extended to the perception of visuotactile stimuli (Hartcher-O'Brien et al., 2008), showing that people tend to miss the tactile stimulus much more often than the visual one on DSS trials with extinction. In this study, Hartcher-O'Brien et al. (2008) first tested their participants in a similar configuration to the audiovisual Colavita experiments done before, that is, with one hand (right) located centrally in front of them. The tactile and visual stimuli were presented from the same spatial location, at the tip of the subject's index finger. In a subsequent experiment, to establish a better correspondence with crossmodal extinction, these authors introduced a spatial dimension to their task, by placing each hand in its corresponding visual field and asking the subjects to fixate centrally. The visual and tactile stimuli could be delivered to either hand, resulting in bimodal trials that were either unilateral (i.e., both stimuli applied to the same hand) or bilateral (i.e., each hand received one of the two stimuli). The latter condition corresponds to the typical experimental set up usually used to test for visuotactile extinction. The results show that, even when both stimuli do not come from the same spatial location but are rather delivered on homologous body parts placed in different hemispaces, the Colavita visual dominance effect still holds, although in lesser proportions.

Following the approach we used to study physiological extinction in the tactile and visual modalities, we recently started to investigate visuotactile extinction-like effects in healthy participants. Tactile stimuli are applied to the tip of either index finger, while visual stimuli are delivered in the space immediately surrounding either hand. A similar intensity titration is used independently for visual and tactile stimuli, on the left and right side, in order for the subject to detect approximately 90% of the single stimuli of each of the four types (2 modalities × 2 sides). Then, during the experiment itself, unilateral visual or tactile stimuli as well as bilateral visuotactile stimulations are randomly delivered while participants are asked to report whether they detected a stimulus, regardless of its modality, on the left, right, or both sides. Preliminary data indicate that visuotactile extinction-like effects can be elicited in healthy individuals using this protocol: participants fail to perceive one of the two stimuli on DSS trials, while perceiving almost perfectly either visual or tactile stimuli delivered in isolation. Here again, there seem not to be a spatial bias in the distribution of their errors.

In sum, it is possible, using well-controlled experimental conditions, to induce extinction-like phenomena in healthy individuals, and thus to establish a physiological model of uni- and multisensory competition. Such a model will allow testing hypotheses drawn from patients' behavior about physiological processes directly, as well as overcoming the potential confounds inherently associated with the study of clinical extinction when it comes to the investigation of the anatomo-functional bases of multisensory perception of competing stimuli. Indeed, although valuable contributions can be made by single case investigations using functional neuroimaging techniques (e.g., Rees et al., 2000; Sarri et al., 2006) as well as group studies using lesion overlapping methods (e.g., Karnath et al., 2003), the neuropsychological approach is limited due to the heterogeneity of lesion sites, sizes, and etiology, and by the plasticity following permanent brain damage. The use of neuroimaging and neurophysiological techniques in neurologically healthy individuals will allow investigating the neural underpinnings of physiological extinction-like phenomena in different modalities without these limitations. After identification of these functional networks, brain stimulation methods such as TMS and tDCS will permit to examine the behavioral disturbances caused by the selective disruption of specific nodes of these networks in the intact brain. This approach will also allow a better comprehension of the anatomo-functional and pathophysiological bases of clinical extinction by comparing these findings to the specific aspects of pathological sensory competition phenomena, such as the ipsilesional spatial bias and the prevalence of extinction after right-hemisphere damage. In sum, thanks to this physiological model of sensory competition, it will be possible to test the validity of current theoretical accounts of extinction, and to better understand how multisensory processes are organized in the intact brain.

In addition to this human physiological model, a refined animal model of the cerebral bases of sensory competition would provide the unique opportunity to bridge the gap between investigations of multisensory processes at the level of single cells and of functional networks, and between studies of the functioning of intact and damaged systems. To this end, establishing a model of physiological extinction-like phenomena in monkeys following the same approach as used in neurologically healthy humans would constitute a very valuable tool in order to refine the framework provided by previous lesion studies in non-human primates.

## TOWARD A NON-HUMAN PRIMATE MODEL OF MULTISENSORY COMPETITION

As suggested in the previous sections, the properties of multimodal neurons recorded in several brain structures including the parietal and premotor cortices in the monkey (Rizzolatti et al., 1981; Graziano and Gross, 1993; Graziano et al., 1994; Duhamel et al., 1998) suggest that these neurons likely constitute crucial nodes in crossmodal interactions and the building of multisensory representations of space. However, typical lesion-based approaches in the monkey have so far been only partially satisfactory.

Cortical ablation studies in monkeys have shown that neglect and extinction can occur following lesions of either the periarcuate frontal (Rizzolatti et al., 1983; Heilman et al., 1995) or the inferior parietal regions (Faugier-Grimaud et al., 1978; Lynch and McLaren, 1989). A first limitation of these studies lies in the fact that the choice of the areas to be lesioned was based on a fair, but putative homology of functional anatomy between humans and monkeys. A second limitation comes from the rapid spontaneous recovery that follows permanent lesions that were used in most studies, leaving only little testing time, sometimes only a few days, which is insufficient to explore the many different aspects of neglect and spatial attention uncovered by neuropsychological investigations in humans (Matelli et al., 1982; Heilman et al., 1995; but see also, e.g., Wardak et al., 2006). Further, this rapid recovery raises questions about the validity of the comparison with human neglect, as it might reflect a true difference in the cerebral organization of attention and space representation mechanisms between the two species, or be related to secondary factors such as lesion location and size, or extent of white matter damage. Third, although testing tasks were appropriate, they were not necessarily sensitive enough.

The emergence of fMRI in monkeys in the recent years provides an invaluable tool to bridge the gap between these findings in monkeys and the research conducted in humans. In particular, one promising perspective is to induce physiological extinction-like phenomena in monkeys within and between sensory modalities using very similar procedures as the one used in humans. The unique advantage of monkey fMRI is that it can be combined with invasive techniques, in particular focal reversible lesions, and thus unequivocally demonstrate which neural network is actually crucial for managing the competition involved in tasks involving the perception of simultaneous sensory stimuli.

By applying the very same task in healthy humans and in trained monkeys to identify the networks involved in competitive sensory perception using fMRI, future work should allow establishing true homologies between humans and monkeys based on functional activations. Functionally identified target regions could then be temporarily and reversibly shut down (using muscimol, see e.g., Wardak et al., 2004, 2006) in order to investigate their causal role in producing behavioral extinction-like effects. In other words, the pitfalls of permanent ablations would be overcome, and the consistency of the results would be ascertained by multiple testing sessions. In addition, using fMRI, one could study the impact of a circumscribed lesion on the activity of the remaining competitive sensory perception network and therefore shed light on the pathophysiology of neglect/extinction. In addition, such a model would undoubtedly provide extremely valuable insight into more theoretical aspects of spatial awareness.

## CONCLUSION

The study of pathological crossmodal extinction has allowed deepening significantly our knowledge about how multisensory processing contributes to human spatial cognition. It has revealed the modular organization of the neural representation of space, with the existence of multiple multisensory maps of different sectors of space likely centered on specific body parts, and allowed to suggest a parallel between behavioral findings in humans and multisensory processes described at the cellular level in non-human primates. It also laid the bases for the development of a model of multisensory competition in neurologically healthy humans. Now, this physiological model can in turn motivate the parallel investigation of competitive multisensory processes in humans and monkeys using complementary techniques in order to complete the framework of the cerebral organization of the perception of multisensory competitive stimuli and of selective spatial attention, in human and non-human primates, from the cellular level to that of functional networks.

## ACKNOWLEDGMENTS

Alessandro Farnè is supported by INSERM, and Martine Meunier by CNRS. This work was funded by the ANR grants no. RPV08085CSA and ANR-O8-BLAN-0068-1, the AVENIR grant no. R05265CS, a scholar award from the James S. McDonnell Foundation, and the NEURODIS Foundation.

## REFERENCES

Becker, E., and Karnath, H. O. (2007). Incidence of visual extinction after left versus right hemisphere stroke. *Stroke* 38, 3172–3174.

Bellas, D. N., Novelly, R. A., Eskenazi, B., and Wasserstein, J. (1988a). The nature of unilateral neglect in the olfactory sensory system. *Neuropsychologia* 26, 45–52.

Bellas, D. N., Novelly, R. A., Eskenazi, B., and Wasserstein, J. (1988b). Unilateral displacement in the olfactory sense: a manifestation of the unilateral neglect syndrome. *Cortex* 24, 267–275.

Bender, M. B. (1952). *Disorders in Perception: With Particular Reference to the Phenomena of Extinction and Displacement*. Springfield, IL: Charles C. Thomas.

Bender, M. B., and Feldman, D. S. (1952). Extinction of taste sensation on double simultaneous stimulation. *Neurology* 2, 195–202.

Bender, M. B., and Teuber, H. L. (1946). Phenomena of fluctuation, extinction and completion in visual perception. *Arch. Neurol. Psychiatry* 55 627–658.

Benton, A. L. (1956). Jacques Loeb and the method of double stimulation. *J. Hist. Med. Allied Sci.* 11, 47–53.

Berlucchi, G., Moro, V., Guerrini, C., and Aglioti, S. (2004). Dissociation between taste and tactile extinction on the tongue after right brain damage. *Neuropsychologia* 42, 1007–1016.

Berti, A., and Frassinetti, F. (2000). When far becomes near: remapping of space by tool use. *J. Cogn. Neurosci.* 12, 415–420.

Bisiach, E., Capitani, E., Luzzatti, C., and Perani, D. (1981). Brain and conscious representation of outside reality. *Neuropsychologia* 19, 543–551.

Bisiach, E., Perani, D., Vallar, G., and Berti, A. (1986). Unilateral neglect: personal and extra-personal. *Neuropsychologia* 24, 759–767.

Bisiach, E., Pizzamiglio, L., Nico, D., and Antonucci, G. (1996). Beyond unilateral neglect. *Brain* 119, 851–857.

Brozzoli, C., Demattè, M. L., Pavani, F., Frassinetti, F., and Farnè, A. (2006). Neglect and extinction: within and between sensory modalities. *Restor. Neurol. Neurosci.* 24, 217–232.

Calvert, G., Spence, C., Stein, B. E., and MITCogNet. (2004). *The Handbook of Multisensory Processes*. Cambridge, MA: MIT press.

Cardinali, L., Brozzoli, C., and Farnè, A. (2009). Peripersonal space and body schema: two labels for the same concept? *Brain Topogr.* 21, 252–260.

Ciçek, M., Gitelman, D., Hurley, R. S., Nobre, A., and Mesulam, M. (2007). Anatomical physiology of spatial extinction. *Cereb. Cortex* 17, 2892–2898.

Cocchini, G., Cubelli, R., Della Sala, S., and Beschin, N. (1999). Neglect without extinction. *Cortex* 35, 285–313.

Colavita, F. B. (1974). Human sensory dominance. *Percept. Psychophys.* 16, 409–412.

Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., and Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nat. Neurosci.* 3, 292–297.

Corbetta, M., Kincade, M., Lewis, C., Snyder, A., and Sapir, A. (2005). Neural basis and recovery of spatial attention deficits in spatial neglect. *Nat. Neurosci.* 8, 1603–1610.

Critchley, M. (1953). *The Parietal Lobes*. New York: Hafner.

Dambeck, N., Sparing, R., Meister, I. G., Wienemann, M., Weidemann, J., Topper, R., and Boroojerdi, B. (2006). Interhemispheric imbalance during visuospatial attention investigated by unilateral and bilateral TMS over human parietal cortices. *Brain Res.* 1072, 194–199.

De Renzi, E., Gentilini, M., and Pattacini, F. (1984). Auditory extinction following hemisphere damage. *Neuropsychologia* 22, 733–744.

di Pellegrino, G., Basso, G., and Frassinetti, F. (1997a). Spatial extinction on double asynchronous stimulation. *Neuropsychologia* 35, 1215–1223.

di Pellegrino, G., Làdavas, E., and Farnè, A. (1997b). Seeing where your hands are. *Nature* 388, 730.

di Pellegrino, G., Basso, G., and Frassinetti, F. (1998). Visual extinction as a spatio-temporal disorder of selective attention. *Neuroreport* 9, 835–839.

di Pellegrino, G., and De Renzi, E. (1995). An experimental investigation on the nature of extinction. *Neuropsychologia* 33, 153–170.

Driver, J., Mattingley, J. B., Rorden, C., and Davis, G. (1997). "Extinction as a paradigm measure of attentional bias and restricted capacity following brain injury," in *Parietal Lobe Contributions to the Orientation in 3D Space*, eds P. Thier and H. O. Karnath (Heidelberg: Springer), 401–430.

Driver, J., and Vuilleumier, P. (2001). Perceptual awareness and its loss in unilateral neglect and extinction. *Cognition* 79, 39–88.

Duhamel, J. R., Colby, C. L., and Goldberg, M. E. (1998). Ventral intraparietal area of the macaque: congruent visual and somatic response properties. *J. Neurophysiol.* 79, 126–136.

Duncan, J. (1984). Selective attention and the organization of visual information. *J. Exp. Psychol. Gen.* 113, 501–517.

Duncan, J. (1996). "Cooperating brain systems in selective perception and action," in *Attention and Performance XVI: Information Integration in Perception and Communication*, eds T. Inui and J. L. McClelland (Cambridge, MA: MIT Press), 549–578.

Ettlinger, G., and Kalsbeck, J. E. (1962). Changes in tactile discrimination and in visual reaching after successive and simultaneous bilateral posterior parietal ablations in the monkey. *J. Neurol. Neurosurg. Psychiatr.* 25 256–268.

Farah, M. J., Wong, A. B., Monheit, M. A., and Morrow, L. A. (1989). Parietal lobe mechanisms of spatial attention: modality-specific or supramodal? *Neuropsychologia* 27, 461–470.

Farnè, A., Brozzoli, C., Làdavas, E., and Ro, T. (2007). "Investigating multisensory spatial cognition through the phenomenon of extinction," in *Attention and Performance* XXII, eds P. Hagard, Y. Rossetti, and M. Kawato (Oxford: Oxford University Press), 183–206.

Farnè, A., Demattè, M. L., and Làdavas, E. (2005a). Neuropsychological evidence of modular organization of the near peripersonal space. *Neurology* 65, 1754–1758.

Farnè, A., Bonifazi, S., and Làdavas, E. (2005b). The role played by tool-use and tool-length on the plastic elongation of peri-hand space: a single case study. *Cogn. Neuropsychol.* 22, 408–418.

Farnè, A., and Làdavas, E. (2000). Dynamic size-change of hand peripersonal space following tool use. *Neuroreport* 11, 1645–1649.

Farnè, A., and Làdavas, E. (2002). Auditory peripersonal space in humans. *J. Cogn. Neurosci.* 14, 1030–1043.

Faugier-Grimaud, S., Frenois, C., and Stein, D. G. (1978). Effects of posterior parietal lesions on visually guided behavior in monkeys. *Neuropsychologia* 16, 151–168.

Fink, G. R., Driver, J., Rorden, C., Baldeweg, T., and Dolan, R. J. (2000). Neural consequences of competing stimuli in both visual hemifields: a physiological basis for visual extinction. *Ann. Neurol.* 47, 440–446.

Geeraerts, S., Lafosse, C., Vandenbussche, E., and Verfaillie, K. (2005). A psychophysical study of visual extinction: ipsilesional distractor interference with contralesional orientation thresholds in visual hemineglect patients. *Neuropsychologia* 43, 530–541.

Gorea, A., and Sagi, D. (2002). Natural extinction: a criterion shift phenomenon. *Vis. Cogn.* 9, 913–936.

Graziano, M. S., and Gross, C. G. (1993). A bimodal map of space: somatosensory receptive fields in the macaque putamen with corresponding visual receptive fields. *Exp. Brain Res.* 97, 96–109.

Graziano, M. S., Hu, X. T., and Gross, C. G. (1997). Visuospatial properties of ventral premotor cortex. *J. Neurophysiol.* 77, 2268–2292.

Graziano, M. S., Reiss, L. A., and Gross, C. G. (1999). A neuronal representation of the location of nearby sounds. *Nature* 397, 428–430.

Graziano, M. S., Yap, G. S., and Gross, C. G. (1994). Coding of visual space by premotor neurons. *Science* 266, 1054–1057.

Halligan, P. W., and Marshall, J. C. (1991). Left neglect for near but not far space in man. *Nature* 350, 498–500.

Hartcher-O'Brien, J., Gallace, A., Krings, B., Koppen, C., and Spence, C. (2008). When vision 'extinguishes' touch in neurologically-normal people: extending the Colavita visual dominance effect. *Exp. Brain Res.* 186, 643–658.

Heilman, K. M., Pandya, D. N., and Geschwind, N. (1970). Trimodal inattention following parietal lobe ablations. *Trans. Am. Neurol. Assoc.* 95, 259–261.

Heilman, K. M., Valenstein, E., Day, A., and Watson, R. (1995). Frontal lobe neglect in monkeys. *Neurology* 45, 1205–1210.

Heilman, K. M., Watson, R. T., and Valenstein, E. (1993). "Neglect and related disorders," in *Clinical Neuropsychology*, 3rd Edn., eds K. M. Heilman and E. Valenstein (New York: Oxford University Press), 279–336.

Heilman K. M., Watson, R. T., and Valenstein, E. (1997). "Neglect: clinical and anatomic aspects," in *Behavioral Neurology and Neuropsychology*, eds T. E. Fineberg and M. J. Farah (New York: McGraw-Hill), 314.

Holmes, N. P., Calvert, G. A., and Spence, C. (2004). Extending or projecting peripersonal space with tools? Multisensory interactions highlight only the distal and proximal ends of tools. *Neurosci. Lett.* 372, 62–67.

Husain, M., Shapiro, K., Martin, J., and Kennard, C. (1997). Abnormal temporal dynamics of visual attention in spatial neglect patients. *Nature* 385, 154–156.

Inhoff, A. W., Rafal, R. D., and Posner, M. J. (1992). Bimodal extinction without cross-modal extinction. *J. Neurol. Neurosurg. Psychiatr.* 55, 36–39.

Iriki, A., Tanaka, M., and Iwamura, Y. (1996). Coding of modified body schema during tool use by macaque postcentral neurones. *Neuroreport* 7, 2325–2330.

Jacquin-Courtois, S., Rode, G., Pavani, F., O'Shea, J., Giard, M. H., Boisson, D., and Rossetti, Y. (2010). Effect of prism adaptation on left dichotic listening deficit in neglect patients: glasses to hear better? *Brain* 133, 895–908.

Jeannerod, M., and Biguer, B. (1987). "The directional coding of reaching movements. A visuomotor conception of spatial neglect," in *Neurophysiological and Neuropsychological Aspects of Spatial Neglect*, Vol. 45 (Amsterdam: North-Holland), 87–113.

Karnath, H. O., Himmelbach, M., and Küker, W. (2003). The cortical substrate of visual extinction. *Neuroreport* 14, 437–442.

Kerkhoff, G. (2001). Spatial hemineglect in humans. *Prog. Neurobiol.* 63, 1–27.

Kinsbourne, M. (1977). Hemi-neglect and hemisphere rivalry. *Adv. Neurol.* 18, 41–49.

Kinsbourne, M. (1987). "Mechanisms of unilateral neglect," in *Neurophysiological and Neuropsychological Aspects of Spatial Neglect*, eds M. Jeannerod (Amsterdam: Elsevier Science Publishers B.V. North-Holland), 69–86.

Kobal, G., Van Toller, S., and Hummel, T. (1989). Is there directional smelling? *Experientia* 45, 130–132.

Làdavas, E., di Pellegrino, G., Farnè, A., and Zeloni, G. (1998a). Neuropsychological evidence of an integrated visuotactile representation of peripersonal space in humans. *J. Cogn. Neurosci.* 10, 581–589.

Làdavas, E., Zeloni, G., and Farnè, A. (1998b). Visual peripersonal space centred on the face in humans. *Brain* 121(Pt 12), 2317–2326.

Làdavas, E., and Farnè, A. (2004). Visuo-tactile representation of near-the-body space. *J. Physiol. Paris* 98, 161–170.

Làdavas, E., Pavani, F., and Farnè, A. (2001). Auditory peripersonal space in humans: a case of auditory-tactile extinction. *Neurocase* 7, 97–103.

Longo, M. R., and Lourenco, S. F. (2006). On the nature of near space: effects of tool use and the transition to far space. *Neuropsychologia* 44, 977–981.

Lynch, J. C., and McLaren, J. W. (1989). Deficits of visual attention and saccadic eye movements after lesions of parietooccipital cortex in monkeys. *J. Neurophysiol.* 61, 74–90.

Maravita, A., and Iriki, A. (2004). Tools for the body (schema). *Trends Cogn. Sci.* 8, 79–86.

Maravita, A., Spence, C., Kennett, S., and Driver, J. (2002). Tool-use changes multimodal spatial interactions between vision and touch in normal humans. *Cognition* 83, B25–B34.

Marcel, A., Postma, P., Gillmeister, H., Cox, S., Rorden, C., Nimmo-Smith, I., and Mackintosh, B. (2004). Migration and fusion of tactile sensation – premorbid susceptibility to allochiria, neglect and extinction? *Neuropsychologia* 42, 1749–1767.

Mark, V. W., Kooistra, C. A., and Heilman, K. M. (1988). Hemispatial neglect affected by non-neglected stimuli. *Neurology* 38, 1207–1211.

Matelli, M., Pavesi, G., and Rizzolatti, G. (1982). Unilateral lesions of the periarcuate cortex in the monkey. *Boll. Soc. Ital. Biol. Sper.* 58, 66–70.

Matthias, E., Bublak, P., Costa, A., Müller, H. J., Schneider, W. X., and Finke, K. (2009). Attentional and sensory effects of lowered levels of intrinsic alertness. *Neuropsychologia* 47, 3255–3264.

Mattingley, J. B., Driver, J., Beschin, N., and Robertson, I. H. (1997). Attentional competition between modalities: extinction between touch and vision after right hemisphere damage. *Neuropsychologia* 35, 867–880.

Meador, K. J., Ray, P. G., Day, L. J., and Loring, D. W. (2001). Relationship of extinction to perceptual thresholds for single stimuli. *Neurology* 56, 1044–1047.

Meister, I. G., Wienemann, M., Buelte, D., Grünewald, C., Sparing, R., Dambeck, N., and Boroojerdi, B. (2006). Hemiextinction induced by transcranial magnetic stimulation over the right temporo-parietal junction. *Neuroscience* 142, 119–123.

Mesulam, M. M. (1981). A cortical network for directed attention and unilateral neglect. *Ann. Neurol.* 10, 309–325.

Oliveri, M., Bisiach, E., Brighina, F., Piazza, A., La Bua, V., Buffa, D., and Fierro, B. (2001). rTMS of the unaffected hemisphere transiently reduces contralesional visuospatial hemineglect. *Neurology* 57, 1338–1340.

Oliveri, M., and Caltagirone, C. (2006). Suppression of extinction with TMS in humans: from healthy controls to patients. *Behav. Neurol.* 17, 163–167.

Oliveri, M., Rossini, P. M., Traversa, R., Cicinelli, P., Filippi, M. M., Pasqualetti, P., Tomaiuolo, F., and Caltagirone, C. (1999). Left frontal transcranial magnetic stimulation reduces contralesional extinction in patients with unilateral right brain damage. *Brain* 122(Pt 9), 1731–1739.

Ortigue, S., Jabaudon, D., Landis, T., Michel, C. M., Maravita, A., and Blanke, O. (2005). Preattentive interference between touch and audition: a case study on multisensory alloesthesia. *Neuroreport* 16, 865–868.

Posner, M. I., Cohen, Y., and Rafal, R. D. (1982). Neural systems control of spatial orienting. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 298, 187–198.

Posner, M. I., Inhoff, A. W., Friedrich, F. J., and Cohen, A. (1987a). Isolating attentional systems: a cognitive-anatomical analysis. *Psychobiology* 15, 107–121.

Posner, M. I., Walker, J. A., Fried, F. A., and Rafal, R. D. (1987b). How do the parietal lobes direct covert attention? *Neuropsychologia* 25, 135–145.

Posner, M. I., Walker, J. A., Friedrich, F. J., and Rafal, R. D. (1984). Effects of parietal injury on covert orienting of attention. *J. Neurosci.* 4, 1863–1874.

Rees, G., Wojciulik, E., Clarke, K., Husain, M., Frith, C., and Driver, J. (2000). Unconscious activation of visual cortex in the damaged right hemisphere of a parietal patient with extinction. *Brain* 123(Pt 8), 1624–1633.

Riddoch, M. J., Rappaport, S. J., and Humphreys, G. W. (2009). Extinction: a window into attentional competition. *Prog. Brain Res.* 176, 149–159.

Rizzolatti, G., Matelli, M., and Pavesi, G. (1983). Deficits in attention and movement following the removal of postarcuate (area 6) and prearcuate (area 8) cortex in macaque monkeys. *Brain* 106(Pt 3), 655–673.

Rizzolatti, G., Scandolara, C., Matelli, M., and Gentilucci, M. (1981). Afferent properties of periarcuate neurons in macaque monkeys. II. Visual responses. *Behav. Brain Res.* 2, 147–163.

Rorden, C., Jelsone, L., Simon-Dack, S., Baylis, L. L., and Baylis, G. C. (2009). Visual extinction: the effect of temporal and spatial bias. *Neuropsychologia* 47, 321–329.

Rorden, C., Mattingley, J. B., Karnath, H. O., and Driver, J. (1997). Visual extinction and prior entry: impaired perception of temporal order with intact motion perception after unilateral parietal damage. *Neuropsychologia* 35, 421–433.

Sarri, M., Blankenburg, F., and Driver, J. (2006). Neural correlates of crossmodal visual-tactile extinction and of tactile awareness revealed by fMRI in a right-hemisphere stroke patient. *Neuropsychologia* 44, 2398–2410.

Serino, A., Pizzoferrato, F., and Làdavas, E. (2008). Viewing a face (especially one's own face) being touched enhances tactile perception on the face. *Psychol. Sci.* 19, 434–438.

Sparing, R., Thimm, M., Hesse, M. D., Kust, J., Karbe, H., and Fink, G. R. (2009). Bidirectional alterations of interhemispheric parietal balance by non-invasive cortical stimulation. *Brain* 132, 3011–3020.

Spence, C. (2009). Explaining the Colavita visual dominance effect. *Prog. Brain Res.* 176, 245–258.

Spence, C., and Driver, J. (2004). *Crossmodal Space and Crossmodal Attention*. Oxford: Oxford University Press.

Stein, B. E., and Arigbede, M. O. (1972). Unimodal and multimodal response properties of neurons in the cat's superior colliculus. *Exp. Neurol.* 36, 179–196.

Vallar, G., Rusconi, M. L., Bignamini, L., Geminiani, G., and Perani, D. (1994). Anatomical correlates of visual and tactile extinction in humans: a clinical CT scan study. *J. Neurol. Neurosurg. Psychiatry* 57, 464–470.

Vuilleumier, P., Valenza, N., Mayer, E., Reverdin, A., and Landis, T. (1998). Near and far visual space in unilateral neglect. *Ann. Neurol.* 43, 406–410.

Wardak, C., Ibos, G., Duhamel, J. R., and Olivier, E. (2006). Contribution of the monkey frontal eye field to covert visual attention. *J. Neurosci.* 26, 4228–4235.

Wardak, C., Olivier, E., and Duhamel, J. R. (2004). A deficit in covert attention after parietal cortex inactivation in the monkey. *Neuron* 42, 501–508.

Watson, R. T., Valenstein, E., Day, A., and Heilman, K. M. (1994). Posterior neocortical systems subserving awareness and neglect. Neglect associated with superior temporal sulcus but not area 7 lesions. *Arch. Neurol.* 51, 1014–1021.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 21 January 2011; accepted: 26 April 2011; published online: 06 May 2011.*

*Citation: Jacobs S, Brozzoli C, Hadj-Bouziane F, Meunier M and Farnè A (2011) Studying multisensory processing and its role in the representation of space through pathological and physiological crossmodal extinction. Front. Psychology **2**:89. doi: 10.3389/fpsyg.2011.00089*

*This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.*

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 23 May 2011
doi: 10.3389/fpsyg.2011.00108

# Racism and the empathy for pain on our skin

**Matteo Forgiarini*, Marcello Gallucci* and Angelo Maravita**

*Department of Psychology, University of Milano-Bicocca, Milan, Italy*

**Edited by:**
*Nadia Bolognini, University of Milano-Bicocca, Italy*

**Reviewed by:**
*Nadia Bolognini, University of Milano-Bicocca, Italy*
*Marco Perugini, University of Milano-Bicocca, Italy*

***Correspondence:**
*Marcello Gallucci and Matteo Forgiarini, Faculty of Psychology, Milan Bicocca University, Piazza dell'Ateneo Nuovo, 1-20126 Milano, Italy.*
*e-mail: marcello.gallucci@unimib.it;*
*e-mail: matteo.forgiarini@gmail.com*

Empathy is a critical function regulating human social life. In particular, empathy for pain is a source of deep emotional feelings and a strong trigger of pro-social behavior. We investigated the existence of a racial bias in the emotional reaction to other people's pain and its link with implicit racist biases. Measuring participants' physiological arousal, we found that Caucasian observers reacted to pain suffered by African people significantly less than to pain of Caucasian people. The reduced reaction to the pain of African individuals was also correlated with the observers' individual implicit race bias. The role of others' race in moderating empathic reactions is a crucial clue for understanding to what extent social interactions, and possibly integration, may be influenced by deeply rooted automatic and uncontrollable responses.

**Keywords: racism, empathy, prejudice, pain, skin conductance**

## INTRODUCTION

Empathy is the ability to understand and vicariously share the feelings and thoughts of other people (De Vignemont and Singer, 2006). Empathic feelings are fundamental for humans in social and interpersonal life because they enable human beings to tune their mental states to their social environment as well as to understand others' intentions, actions, and behaviors. One of the main sources of empathic feelings is the pain experienced by other human beings, and empathy for others' pain, in turn, regulates behavior among individuals and social groups. Although pain has been considered an intimate and private feeling, experimental data indicate that when people witness or imagine the pain of another person, they map the others' pain onto their brain using the same network activated during firsthand experience of pain, as if they were vicariously experiencing the observed pain (Hutchison et al., 1999; Carr et al., 2003; Wicker et al., 2003; Morrison et al., 2004; Singer et al., 2004; Bufalari et al., 2007; Lamm et al., 2007; Cheng et al., 2008). Furthermore feeling other people's suffering triggers pro-social behavior (Batson et al., 2002), promotes helping and encourages cooperation (Batson et al., 1997a). Conversely, lack of empathy for the pain of other human beings may lead to violence, abuse, and deterioration of interpersonal and intergroup relationships (Batson et al., 2002).

Empathic reactions to pain involve different layers of cognitive processing, with a predominant role played by automatic and implicit processes. Recent neurophysiological findings (Singer et al., 2004) have documented specific neuropsychological activations of the affective but not sensory components of the brain (the pain matrix in particular), leading to fast and automatic responses to the pain of others. Similarly, the vision of a needle penetrating the hand reduces the muscular motor response in the observer compatible with the locus of injection in the target person (Avenanti et al., 2005). Immediate empathic reactions, however, are deeply affected by social cues and individual differences. We now know that functional activity related to empathy reactions to others' feelings is affected by similarity between the witness and the person in pain (Krebs, 1975; Preston and de Waal, 2001; Lamm et al., 2010), by previous experience in the same situation (Cheng et al., 2007), by observer's personality (Avenanti et al., 2009), by social relationship with the target (Singer et al., 2006), by familiarity for the target (Cialdini et al., 1997), gender (Eisenberg and Carlo, 1995; Baron-Cohen and Wheelwright, 2004), and age (Eisenberg and Morris, 2001). Most relevant for the present work, initial findings seem to suggest that empathy for others pain can be modulated by ingroup/outgroup social categorization based on race differences between the target and the onlooker (Cosmides et al., 2003; Xu et al., 2009; Avenanti et al., 2010).

If immediate empathic reactions differ depending on the race of the target person, this modulation may be rooted in personal characteristics of the person experiencing empathy, allowing for individual differences in the strength of the link between race and empathy. A weaker reaction to pain of other race members may be a consequence of personal prejudices and, more generally, racist attitudes toward the outgroup member. It is unlikely, however, that immediate physiological reactions are modulated by explicit and elaborated cognitive processes. We reckon that the possible link between racist attitudes and weaker empathy reactions for others' pain should operate at an implicit level, where prejudices and racism manifest themselves through fast and unconscious associations between negative evaluations and the target race (Greenwald et al., 1998). Along this line of reasoning, a recent independent TMS study (Avenanti et al., 2010) using a Blacks/Whites measure of implicit racist attitude (Greenwald et al., 1998), showed that despite the lack of explicit racial bias in the sample, participants with higher implicit ingroup preference presented greater corticospinal reactivity to ingroup models over outgroup models' pain.

The present research is aimed at providing experimental evidence that automatic, physiological reactions to other people's pain strongly depends on the race of the person in pain, such that pain received by members of other racial groups elicits a much weaker reaction compared with the pain suffered by members of the same group. By presenting participants with a series of video clips, in two experiments we tested whether the reaction to pain of Caucasian (Italian) observers was influenced by the race (Caucasian, Asian, or African) of the person in pain. In the second study we replicate this finding and show that the moderation of empathy is correlated with the individual implicit racial biases.

Empathic reactions were inferred by the skin conductance responses (SCR; Purves et al., 2008) to observed video stimuli showing human subjects experiencing either harmless or painful somatosensory stimuli. The third-person exposure to pain activates a brain network called "pain matrix" (Peyron et al., 1999, 2000; Derbyshire, 2000) which includes the anterior cingulate cortex. Activity in the ACC is known to trigger variation in the skin conductance (SC; Purves et al., 2008), i.e., SC increases as a physiologic and autonomic response to someone else's pain (Krebs, 1975; Levenson and Ruef, 1992; Morrison et al., 2004; Rae Westbury and Neumann, 2008). The triggering role of the ACC on SC is likely to have been selected to facilitate coping and adaptive responses (Devinsky et al., 1995; Ledowski et al., 2006). Furthermore, SC is considered one of the most reliable predictors of accurately assessed negative emotions in others, as it is associated with the emotional response rather than the mere pain recognition (Levenson and Ruef, 1992).

For the above reasons, the SCR was chosen as a likely marker of the automatic, emphatic response to the vision of painful stimuli inflicted in another person.

## EXPERIMENT 1

### MATERIALS AND METHODS

Stimuli were video clips showing a person subject to a painful stimulus or a harmless one. Each video started with a frame depicting a face of a female or a male actor holding a neutral expression. Subsequently the camera zoomed in on the actor's hand which was touched by the experimenter alternatively by an eraser (harmless stimulus) or by a needle (painful stimulus). A total of 12 video clips were presented to each participant, featuring six different actors: Two Caucasian, two African, and two Asian actors. For each race, a female and a male actor was used, each actor subject to a painful stimulus and a harmless one.

In both experiments, participants sat in front of a computer monitor (Acer aspire 1360, monitor 15.4″ TFT LCD) where the experimental stimuli were displayed. The distance between the monitor and participant's face was almost 70 cm. Prior to stimulus presentation, two electrodes were applied on the forefinger and ring-finger of participant's left hand in order to record the SCR. Participants were asked to relax, and carefully watch the stimuli presented on the monitor. During SCR recording, participants were listening to white noise with headphones in order to cover external auditory stimuli. The videos order was completely randomized. The experimenter, blind to stimuli presentation, started each video after visually checking that the online SCR was returned to a baseline level, in order to avoid response overlaps to consecutive stimuli. Following this procedure the inter stimulus interval was 15 s (range 10–20) across all participants. Each experimental session lasted almost 45 min, including behavioral and psychophysiological data gathering. All participants gave their consents to physiological recording and display of videos prior to the experiments. After the experiments ended, participants were fully debriefed regarding the nature of the stimuli and aim of the study.

Skin conductance responses was measured while the observers viewed the video stimuli. The difference between the SCR subsequent to a painful stimulus and the SCR subsequent to a harmless stimulus was taken as a measure of empathy for pain (hereinafter empathic index, EI). Physiological data collection was performed using The UFI model 2701 BioDerm(R) SC meter. It is a stand-alone instrument which measures skin conductance level (SCL) and SCR. Skin conductance is measured using an Ag–AgCl electrode pair with the constant voltage (0.5 V) method. The SCR were recorded as the phase component of the SC activity, with a 10-Hz rate. Microsiemens (μs) are the measurement units. For all the participants, the SCR recordings were synchronized with the first video frame presented. The average response within a time window of 6.5 s post-stimulus was used as the observed variable for the analysis[1]. This interval was chosen based on the relevant literature (Purves et al., 2008) as the most appropriate and included the whole variation of SC following the stimuli. The SCR data analysis was performed using the SAS General Linear Model procedure. Unless otherwise specified, all the results discussed were obtained with a least squares repeated-measures ANOVA. The estimated means for the stimulus × race interaction at different levels (1 SD below and 1 SD above the mean) of the continuous implicit association test (IAT) variable were obtained with an equivalent model estimated with the PROC MIXED SAS procedure (i.e., SAS procedure commonly used to estimate mixed effects linear models).

Ninety students of Milano-Bicocca University have been recruited. Three participants were excluded for problems in data saving; 5 participants were excluded because of uncooperative behavior during the experiment; 17 participants were excluded due to technical problems during the experiment. Out of the 65 remaining participants, four were excluded as outliers: SCR scores exceeding 2 SD from the overall average. A total of 61 participants (29 female) were therefore included in the analyses.

Participants were subjected to a 2 (stimuli: harmless and painful) × 3 (races of the target person: African, Caucasian, Asian) × 2 (blocks: first and second experimental block) × 2 (target gender: male and female) repeated-measure factorial design.

### RESULTS

Participants showed an overall significant EI: reactions to painful stimuli were significantly greater than reactions to harmless stimuli [$F(1,59) = 40.85$, $P < 0.001$].

Crucially, the race of the actor experiencing the painful stimulus significantly moderated the EI [$F(2,118) = 3.6$, $P = 0.03$]. Although experimental participants showed a significant EI for Caucasians [$F(1,59) = 29.57$, $P < 0.001$, $\eta^2 = 0.333$], for Africans [$F(1,59) = 7.52$, $P = 0.008$, $\eta^2 = 0.113$], and for Asian images [$F(1,59) = 16.99$, $P < 0.001$, $\eta^2 = 0.223$], the empathic reaction for the Caucasians was significantly greater than that for the Africans [$F(1,59) = 7.87$, $P = 0.006$, $\eta^2 = 0.117$; **Figure 1**]. Critically, there was no racial effect on the reaction to the harmless stimuli [$F(2,118) = 0.09$, $P = 0.91$], a significant moderating effect of target person race was found on the reaction to painful stimuli [$F(2,118) = 5.09$, $P = 0.007$]: Reactions to Caucasians painful stimuli were significantly greater than for Africans [$t(118) = 2.91$, $P = 0.004$] but not than for Asian targets [$t(118) = 1.72$, $P = 0.08$].

---

[1] This type of analysis is consistent with technical manual published by the producer of the UFI BioDerm System (http://www.ufiservingscience.com/).

As expected, the SCR responses significantly varied during the time for the painful stimuli, but not for the harmless stimuli. Specifically, the reactions to painful stimuli significantly reduced over time [block effect: *F*(1,59) = 44.58, *P* < 0.0001] and the reactions to harmless stimuli were constant during the experiments [block effect: *F*(1,59) = 0.15, *P* = 0.70]. No other effect was statistically significant.

## EXPERIMENT 2: IMPLICIT ATTITUDES AND EMPATHY FOR PAIN

### MATERIALS AND METHODS

The second experiment aimed at linking the empathic racial bias with the implicit racial prejudice. In addition to using the same paradigm used in Experiment 1, the experiment required participants to complete a race (Caucasians/Africans) IAT (Greenwald et al., 1998) and a Trait Empathy Scale (Mehrabian and Epstein, 1972). Specifically, we assessed to what extent individual differences in the implicit racial prejudice correlate with the difference between the EI for Caucasians with respect to EI for Africans.

The IAT provides access to deep cognitive domains that is not reached by self-report measures (Greenwald et al., 1998). In the version used in this study, it provides a measure of implicit differential evaluation of Caucasian and African races. The IAT is based on participants' reaction times on a computer-based categorization task. The IAT assesses the association between two classes of stimuli by measuring differences in the response speed that participants show in the same task with exemplars from two categories. The task we used rates the association strength between positive and negative concepts with Caucasian and African races. On each trial of the race IAT we used, participants categorized a stimulus from one of four the categories: a photo of a Caucasian man, a photo a African man, a positive word (Joy, love, peace, wonderful pleasure, friend), or a negative word (agony, terrible, awful, bad, evil, war). In one block of trials, positive words required the same behavioral response as photos of Caucasian men. In another block of trials, positive words required the same response as photos of African men. IAT data were coded in the direction of association between positive words and Caucasian targets, i.e., as the difference in mean response latency to trials in positive word-African targets block minus trials in the positive words-Caucasian targets block. Then, higher scores reflect strong associations between positive concepts and Caucasian race as well as strong association between negative concepts and African race. The fourth and seventh blocks consisted of 40 trials, all of the other blocks consisted of 20 trials. IAT scores reported in the text and used in the analysis are D scores. We calculated D scores using Nosek's SAS macro based on Greenwald et al. (2003). Individual IAT scores were used in the mixed model as a continuous independent variable. The simple slopes analysis was conducted to estimate the experimental effects at specific values of the IAT score. This was obtained by centering the IAT scores to 1 SD above the mean (and subsequently at 1 SD below) before entering the IAT score in the model (Aiken and West, 1991).

The experiment included 60 students of Milano-Bicocca University. Two participants were excluded for problems in data saving; five participants were excluded due to technical problems during the experiment. Out of these 53 participants, 6 were excluded as outliers (SCR scores exceeding 2 SD from overall average). A total of 47 participants (24 female) were therefore included in the analyses.

### RESULTS

Results replicated the overall stronger reaction to painful than to harmless stimuli [*F*(1,45) = 36.63, *P* < 0.001]. Target race significantly moderated the EI [*F*(2,90) = 4.26, *P* = 0.01; **Figure 2**]. The EI was significant for Caucasian [*F*(1,45) = 23.85, *P* < 0.0001, $\eta^2$ = 0.346] and Asian [*F*(1,45) = 13.9, *P* = 0.0005, $\eta^2$ = 0.225] but not for African actors [*F*(1,45) = 1.36, *P* = 0.24, $\eta^2$ = 0.029]. Furthermore, the EI was statistically lower for Africans than for Caucasians [*F*(1,45) = 6.64, *P* = 0.01] targets (**Figure 2**). As for Experiment 1, we observed no race differences in reactions for harmless stimuli [*F*(2,90) = 0.1, *P* = 0.9] and a race significant effect on the reaction for other people's pain [*F*(2,90) = 7.55, *P* < 0.001]. The SCR and the IAT scores were analyzed together. In the IAT used, greater scores indicate faster associations of "Caucasians" with positive concepts and "Africans" with negative concepts, thus a

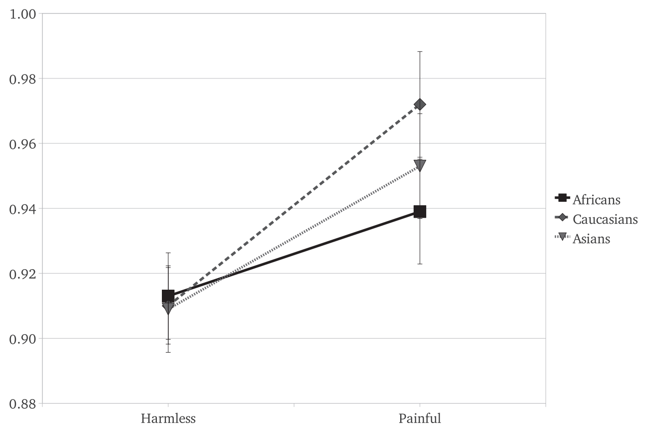

**FIGURE 1 | Experiment 1: mean SCR and standard errors as a function of stimulus type and actor's race.** Responses to pain were always greater than those to the harmless stimuli.

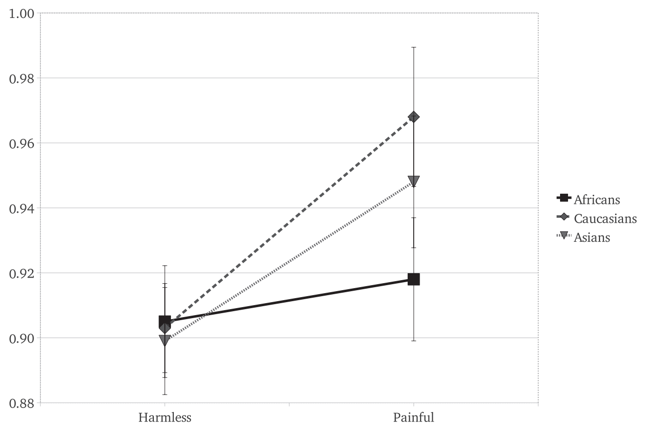

**FIGURE 2 | Experiment 2: SCR means as a function of stimulus type and the race of the person in pain.**

stronger racial bias against "Africans." The IAT scores were included in a GLM comprising the SCR values elicited only by Africans and Caucasians stimuli (the two races included in the IAT). The IAT scores were included in the model as a continuous independent variable. As typically found with the race IAT, Caucasian observers more strongly associated negative stereotypes with Africans than with Caucasians, [$F(1,46) = 34.45$, $P < 0.001$]. Most importantly, the strength of the implicit race bias correlates with the reduced empathy for Africans' pain. We found that the IAT scores of the observers significantly predict the moderating effect of race on the reaction for pain [$F(1,43) = 4.52$, $P = 0.03$]. Simple slope analysis (Aiken and West, 1991) revealed that the greater the participant racial bias, the greater the difference between the empathic responses toward Caucasians with respect to Africans (**Figure 3**). Data show that participants with low race bias (1 SD below sample average) are not affected by the race moderating effect on the empathic responses to actors' pain (**Figure 3B**). Participants EI is significantly greater than zero [$F(1,45) = 5.22$, $P = 0.02$] but it's not moderated by the race of the person in pain [$F(1,45) = 0.14$, $P = 0.70$] and there are no differences in the overall reactions for Caucasians and Africans [$F(1,45) = 1.28$, $P = 0.25$]. On the other hand, data show that for participants with an high race bias (IAT score 1 SD above the sample average) EI is significant greater than zero [$F(1,45) = 14.52$, $P = 0.0001$] and the race of the person in pain significantly moderates the empathic reactions [$F(1,45) = 13.29$, $P = 0.0003$; cf. **Figure 3A**].

The BEES empathy scale completed by participants had no significant effects on the SCR EI for any of three races we tested. Even though the BEES seems to account for the empathy related brain areas activation (Singer et al., 2004), this empathy scale seems to failed in prediction of SCR empathy related activations (Rae Westbury and Neumann, 2008).

As regards effects over time, reactions to painful stimuli significantly reduced over time [block effect: $F(1,45) = 8.08$, $P = 0.006$] and the reactions to harmless stimuli were constant during the experiments [block effect: $F(1,45) = 1.67$, $P = 0.20$]. These results replicated Experiment 1 results, suggesting that participants' stimuli perception were reliable and precise during the entire experiment. No other effect was significant.

In order to rule out possible alternative explanations, we performed the analysis of SCR baseline values immediately before participants' empathic reactions. The mean of SCR values during the 600-ms pre-stimulus was calculated. As expected, in Experiments 1 and 2 the full model revealed no relevant effects for all the experimental factors on the pre-stimulus SCR. These findings rule out the possibility that the observed responses associated with the painful and harmless stimuli were due to stochastic effects prior to stimulus presentation.

## EXPERIMENTS 1 AND 2: GENERAL DISCUSSION

Taken together our findings demonstrate a clear pattern of responses to pain: the extent to which Caucasian observers share the pain experience of other people is affected by the race of the person in pain (**Figure 4A**). Before the stimulus onset, the SCR values show stochastic variations. After observing a painful stimulus administered to the target person, participants' SCR values increase more for Caucasian targets than for target people of the other races, and the least for African targets.

This race moderation pattern was not present in the reactions to harmless stimuli (**Figure 4B**). During the video stimuli perception, before and after the stimulus onset, participants' responses are not affected by the race of the target people.

Given the link between automatic SCR response and emotional response to an observed situation, our data suggest that implicit and uncontrolled cognitive mechanisms lead Caucasians to reduce the automatic sharing of pain experience with African conspecifics at an automatic, early level of stimulus processing. These data concur with studies demonstrating deep connections between implicit social evaluations and neurological responses of the central nervous system. Research in the field of social neuroscience shows that in white participants the strength of amygdale activation to unfamiliar black vs. white faces is correlated with implicit but not explicit race bias measures (Phelps et al., 2000) and is related to different level of PFC activity (Cunningham et al., 2004).

The correlation between the empathic reactions as measured with the SCR and the IAT scores rules out alternative explanations of the effect based on some low-level perceptual features of the video stimuli presented. In fact, perceptual artifacts should have a general reduction of response to African actors for both harmless and painful stimuli,

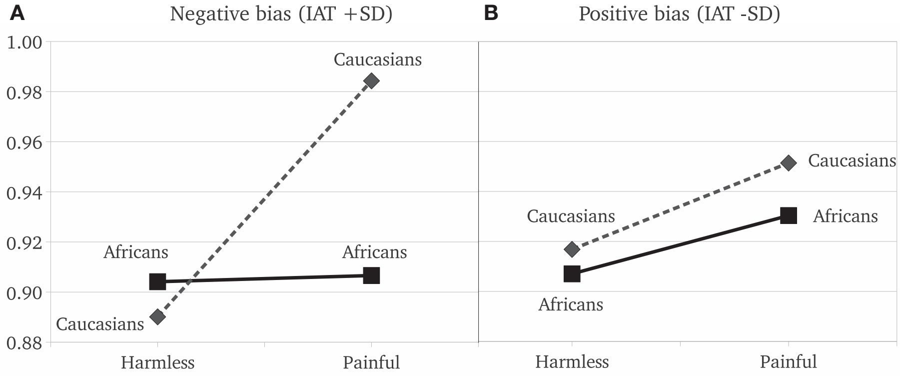

**FIGURE 3 | Experiment 2: simple slope analysis.** Estimated SCR means as a function of stimulus type and the race of the person computed at two different levels of IAT scores: **(A)** estimation for strong negative bias for Blacks (1 SD above average IAT score); **(B)** estimation for positive bias for Blacks (1 SD below average IAT score).

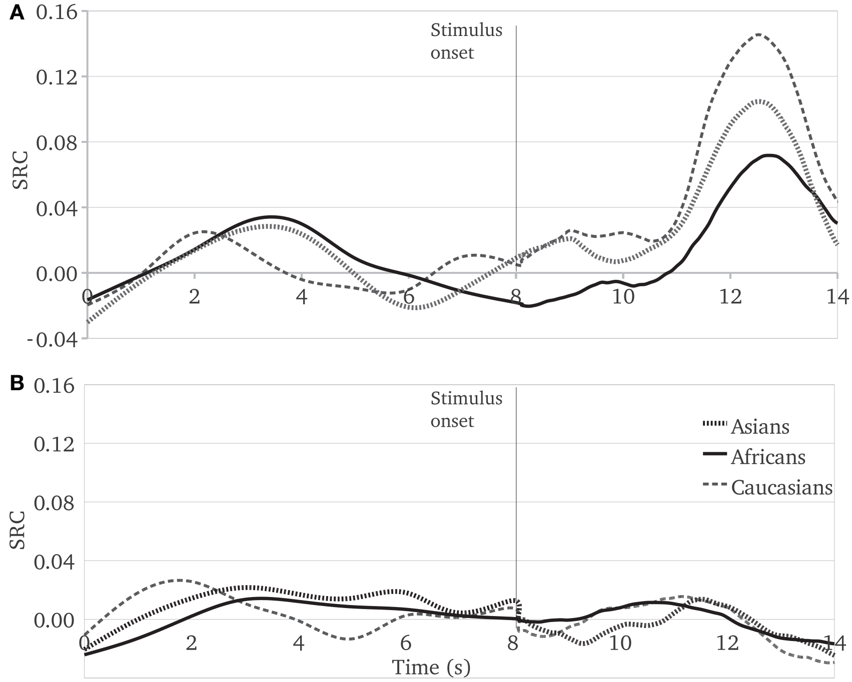

**FIGURE 4 | Experiments 1 and 2, SCR mean values as a function of time and race of target people. (A)** Reactions to painful stimuli, **(B)** Reactions to harmless stimuli.

whereas the moderation due to race is specifically found for painful stimuli[2]. As our data did not show any gender effect, mere similarity between actors and observers could not account for our results.

Interestingly, these data do not support the outgroup antipathy hypothesis (Brown et al., 2006) as they do not indicate increased affective reactions to stimuli of outgroup members in general. Moreover, the ingroup empathy hypothesis (Brown et al., 2006) does not seem to account for our effects either. A mere ingroup–outgroup categorization should lead to a significant reduction of empathy for Africans as well as Asian actors. Instead, the EI to Africans was lower than the one for Caucasians and Asian targets in both experiments. Furthermore, although one of the most pervasive categorizations in human society is gender, we never observed an interaction between the subject's gender and stimulus gender in the empathic responses (in both of the Experiments 1 and 2, *P*s > 0.63). Interestingly findings regarding a differential reaction to animals in pain (Rae Westbury and Neumann, 2008) suggest that empathic feelings in humans are moderated by the perceived phylogenetic similarity between the observer and the suffering animal. In a similar vein, data gathered in the two studies, seem to indicate that the closer the phenotypic aspect of the actor and the observer, the stronger the psychophysiologic empathic response to pain. Consistently with our data Batson et al. (1997b) found that university group membership (i.e., shared or unshared) had no impact on empathetic induction. Conversely, Johnson et al. (2002) showed that the defendant race clearly affected the empathetic induction on a subsequent juror decision-making task. One reason for the divergent findings might involve the differential nature of the groups studied. One obvious explanation is that race might be more relevant or salient than university membership (Cunningham, 1986; Krebs, 1991). In a similar vein, Cosmides et al. (2003) noted that racial group membership defines coalitions and alliances during evolution and thus results in strong modulation of the neural substrates of emotional components of empathy.

Thus, previous research and our findings suggest that relevant group membership might play a significant role in empathetic induction. But how can people rate differences of race? And, what does really means being members of two different races? Since it is very difficult to directly measure the degree of genetic similarity with others, Krebs (1991) has suggested that one relies on discernible cues to make such judgments. Clearly, racial indicators (e.g., skin color, hair texture) would seem to qualify as powerful cues of kinship and genetic similarity. Furthermore recent findings showed that racial biases affect clinical pain management: Pletcher et al. (2008) provided evidence that physicians withhold opioid treatment from Hispanic, Black, and Asian patients compared to White patients, despite similar pain severity. The authors also noted this therapeutic disparity cannot be attributed to patient histories of alcohol and drug abuse as disproportionate treatment was most apparent in patients under the age of 12.

Moreover, our data support the idea that racial groups different from the perceiver could elicit a weaker sense of familiarity than a more similar conspecific. *Dehumanization Theory* (Fiske

---

[2] Although our results are clear and in line with the theoretical expectations, further research is needed to understand the extent to which the effects we found could be replicated using different sets of stimuli and different physiological markers.

et al., 2002) posits that some specific social groups are perceived as less human as they activate non-exclusively social emotions. The so called Stereotyped Content Model predicts that perceived outgroups' warmth and competence affect what emotions outgroups elicit in perceivers. Our data are consistent with this model and findings we reported seem to place race beside warmth and competence as a predictor of shared emotions. According to Singer's theory phenotypically distant outgroups may even elicit different patterns of brain activation, with particular regard to the medial PFC (Harris and Fiske, 2006), a region of cortex implicated in social cognition. This region responds to faces of people belonging to all social groups except extreme outgroups who activate, instead, a pattern consistent with disgust. Different degrees of dehumanization may therefore account for our findings that Caucasians could perceive Asians and Africans on different levels of humanization. Our findings indicate that even if we are not aware, our body and our mind use internalized knowledge to address reactions and activities they engage to deal with social and physical world. Future research may strengthen the link between *Dehumanization Theory* and differential empathic reactions to pain.

Clearly, relationship between race and empathy has many significant implications for real-world issues. Would racial differences decrease the probability that a Black teacher could empathize with a White student? Would racial similarity increase the probability that a White supervisor could empathize with a White subordinate? How a manager choose the person who has to be fired between a Black and a White worker? Although the teacher, the supervisor and the manager might make an effort to empathize, they are likely to be biased and experience less empathy for Blacks involved in their choices.

Our findings suggest that the attitude toward other races may involve not only the overt self-report of the observer concerning attitudes about race but also their deep automatic and physiological reactions. These differential reactions may be elicited even at a very basic level, such as the reaction to physical pain of others. Such a fundamental racial differentiation, in turn, may bias complex activities and judgments over and beyond human consciousness. A precise assessment of other people's pain, in fact, is a necessary skill in many human activities, from medical decisions, rescue operations, police intervention, policy making and, in extreme circumstances, use of physical force and punishment. When all these activities involve people perceived as belonging to different races, a racial bias may hinder pain assessment with detrimental effects on individuals, groups, and their peaceful relationships.

## ACKNOWLEDGMENTS

This work was supported by funding from the University of Milano-Bicocca (FAR) to Marcello Gallucci and Angelo Maravita. Preliminary data were presented to the General Meeting of the European Association of Experimental Social Psychology, 10–14 June 2008, Opatija, Croatia.

## REFERENCES

Aiken, L. S., and West, S. G. (1991). *Multiple Regression: Testing and Interpreting Interactions.* Thousand Oaks: Sage.

Avenanti, A., Bueti, D., Galati, G., and Aglioti, S. M. (2005). Transcranial magnetic stimulation highlights the sensorimotor side of empathy for pain. *Nat. Neurosci.* 8, 955–960.

Avenanti, A., Minio-Paluello, I., Bufalari, I., and Aglioti, S. M. (2009). The pain of a model in the personality of an onlooker: influence of state-reactivity and personality traits on embodied empathy for pain. *Neuroimage* 44, 275–283.

Avenanti, A., Sirigu, A., and Aglioti, S. M. (2010). Racial bias reduces empathic sensorimotor resonance with other-race pain. *Curr. Biol.* 20, 1018–1022.

Baron-Cohen, S., and Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *J. Autism Dev. Disord.* 34, 163–175.

Batson, C. D., Chang, J., Orr, R., and Rowland, J. (2002). Empathy, attitudes and action: can feeling for a member of a stigmatized group motivate one to help the group. *Pers. Soc. Psychol. Bull.* 28, 1656–1666.

Batson, C. D., Polycarpou, M. P., Harmon-Jones, E., Imhoff, H. J., Mitchener, E. C., Bednar, L. L., Klein, T. R., and Highberger, L. (1997a). Empathy and attitudes: can feeling for a member of a stigmatized group improve feelings toward the group? *J. Pers. Soc. Psychol.* 71, 105–118.

Batson, D., Sager, K., Garst, E., Kang, M., Kostia, R., and Karen, D. (1997b). Is empathy-induced helping due to self-other merging? *J. Pers. Soc. Psychol.* 73, 495–509.

Brown, L. M., Bradley, M. M., and Lang, P. J. (2006). Affective reactions to pictures of ingroup and outgroup members. *Biol. Psychol.* 71, 303–311.

Bufalari, I., Aprile, T., Avenanti, A., Di Russo, F., and Aglioti, S. M. (2007). Empathy for pain and touch in the human somatosensory cortex. *Cereb. Cortex* 17, 2553–2561.

Carr, L., Iacoboni, M., Dubeau, M.-C., Mazziotta, J. C., and Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proc. Natl. Acad. Sci. U.S.A.* 100, 5497–5502.

Cheng, Y., Lin, C.-P., Liu, H.-L., Hsu, Y.-Y., Lim, K.-E., Hung, D., and Decety, J. (2007). Expertise modulates the perception of pain in others. *Curr. Biol.* 17, 1708–1713.

Cheng, Y., Yang, C.-Y., Lin, C.-P., Lee, P.-L., and Decety, J. (2008). The perception of pain in others suppresses somatosensory oscillations: a magnetoencephalography study. *Neuroimage* 40, 1833–1840.

Cialdini, R. B., Brown, S. L., Lewis, B. P., Luce, C., and Neuberg, S. L. (1997). Reinterpreting the empathy-altruism relationship: when one into one equals oneness. *J. Pers. Soc. Psychol.* 73, 481–494.

Cosmides, L., Tooby, J., and Kurzban, R. (2003). Perceptions of race. *Trends Cogn. Sci. (Regul. Ed.)* 7, 173–179.

Cunningham, M. (1986). Levites and brother's keepers: a sociobiological perspective on prosocial behavior. *Humboldt J. Soc. Relat.* 13, 35–67.

Cunningham, W. A., Johnson, M. K., Raye, C. L., Gatenby, J. C., Gore, J. C., and Banaji, M. R. (2004). Separable neural components in the processing of black and white faces. *Psychol. Sci.* 15, 806–813.

De Vignemont, F., and Singer, T. (2006). The empathic brain: how, when and why? *Trends Cogn. Sci.* 10, 435–441.

Derbyshire, S. (2000). Exploring the pain "neuromatrix". *Curr. Pain Headache Rep.* 4, 467–477.

Devinsky, O., Morrell, M. J., and Vogt, B. A. (1995). Contributions of anterior cingulate cortex to behaviour. *Brain* 118, 279–306.

Eisenberg, N., and Carlo, G. (1995). Prosocial development in late adolescence: a longitudinal study. *Child Dev.* 66, 1179–1197.

Eisenberg, N., and Morris, A. S. (2001). The origins and social significance of empathy-related responding. A review of empathy and moral development: implications for caring and justice by M. L. Hoffman. *Soc. Justice Res.* 14, 95–120.

Fiske, S. T., Cuddy, A. J., Glick, P., and Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *J. Pers. Soc. Psychol.* 82, 878–902.

Greenwald, A., McGhee, D. E., and Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: the implicit association test. *J. Pers. Soc. Psychol.* 74, 1464–1480.

Greenwald, A. G., Nosek, B. A., and Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *J. Pers. Soc. Psychol.* 85, 197–216.

Harris, L. T., and Fiske, S. T. (2006). Dehumanizing the lowest of the low: neuroimaging responses to extreme out-groups. *Psychol. Sci.* 17, 847–853.

Hutchison, W. D., Davis, K. D., Lozano, A. M., Tasker, R. R., and Dostrovsky, J. O. (1999). Pain-related neurons in the human cingulate cortex. *Nat. Neurosci.* 2, 403–405.

Johnson, J., Simmons, C., Jordan, A., MacLean, L., Taddei, J., and Thomas,

D. (2002). Rodney King and O. J. revisited: the impact of race and defendant empathy induction on judicial decisions. *J. Appl. Soc. Psychol.* 32, 1208–1223.

Krebs, D. (1975). Empathy and altruism. *J. Pers. Soc. Psychol.,* 32, 1134–1146.

Krebs, D. L. (1991). Altruism and egoism: a false dichotomy. *Psychol. Inq.* 2, 137–139.

Lamm, C., Batson, C. D., and Decety, J. (2007). The neural substrate of human empathy: effects of perspective-taking and cognitive appraisal. *J. Cogn. Neurosci.* 19, 42–58.

Lamm, C., Meltzoff, A. N., and Decety, J. (2010). How do we empathize with someone who is not like us? A functional magnetic resonance imaging study. *J. Cogn. Neurosci.* 22, 362–376.

Ledowski, T., Bromilow, J., Paech, M. J., Storm, H., Hacking, R., and Schug, S. A. (2006). Monitoring of skin conductance to assess postoperative pain intensity. *Br. J. Anaesth.* 97, 862–865.

Levenson, R. W., and Ruef, A. M. (1992). Empathy: a physiological substrate. *J. Pers. Soc. Psychol.* 63, 234–246.

Mehrabian, A., and Epstein, N. (1972). A measure of emotional empathy. *J. Pers.* 40, 525–543.

Morrison, I., Lloyd, D., Di Pellegrino, G., and Roberts, N. (2004). Vicarious responses to pain in anterior cingulate cortex: is empathy a multisensory issue? *Cogn. Affect. Behav. Neurosci.* 4, 270–278.

Peyron, R., Garcia-Larrea, L., Gregoire, M.-C., Costes, N., Convers, P., Lavenne, F., Mauguiere, F., Michel, D., and Laurent, B. (1999). Haemodynamic brain responses to acute pain in humans: sensory and attentional networks. *Brain* 122, 1765–1780.

Peyron, R., Laurent, B., and García-Larrea, L. (2000). Functional imaging of brain responses to pain. A review and meta-analysis (2000). *Neurophysiol. Clin.* 30, 263–288.

Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., and Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *J. Cogn. Neurosci.* 12, 729–738.

Pletcher, M. J., Kertesz, S. G., Kohn, M. A., and Gonzales, R. (2008). Trends in opioid prescribing by race/ethnicity for patients seeking care in US emergency departments. *JAMA* 299, 70–78.

Preston, S. D., and de Waal, F. B. M. (2001). Empathy: its ultimate and proximate bases. *Behav. Brain Sci.* 25, 1–20.

Purves, D., Brannon, E. M., Cabeza, R., Huettel, S. A., LaBar, K. S., Platt, M. L., and Woldorff, M. G. (2008). *Principles of Cognitive Neuroscience.* Sunderland, MA: Sinauer Associates Inc.

Rae Westbury, H., and Neumann, D. L. (2008). Empathy-related responses to moving film stimuli depicting human and non-human animal targets in negative circumstances. *Biol. Psychol.* 78, 66–74.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., and Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science* 303, 1157–1162.

Singer, T., Seymour, B., O'Doherty, J. P., Stephan, K. E., Dolan, R. J., and Frith, C. D. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature* 439, 466–469.

Wicker, B., Keysers, C., Plailly, J., Royet, J.-P., Gallese, V., and Rizzolatti, G. (2003). Both of us disgusted in my insula: the common neural basis of seeing and feeling disgust. *Neuron* 40, 655–664.

Xu, X., Zuo, X., Wang, X., and Han, S. (2009). Do you feel my pain? Racial group membership modulates empathic neural responses. *J. Neurosci.* 29, 8525–8529.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 26 January 2011; accepted: 12 May 2011; published online: 23 May 2011.*
*Citation: Forgiarini M, Gallucci M and Maravita A (2011) Racism and the empathy for pain on our skin. Front. Psychology* ***2****:108. doi: 10.3389/fpsyg.2011.00108*
*This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.*

frontiers in
PSYCHOLOGY

REVIEW ARTICLE
published: 23 December 2011
doi: 10.3389/fpsyg.2011.00383

# Multi-sensory and sensorimotor foundation of bodily self-consciousness – an interdisciplinary approach

***Silvio Ionta*[1]*, Roger Gassert*[2] *and Olaf Blanke*[1,3]****

[1] *Laboratory of Cognitive Neuroscience, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland*
[2] *Rehabilitation Engineering Laboratory, Eidgenössische Technische Hochschule Zürich, Zurich, Switzerland*
[3] *Department of Neurology, University Hospital, Geneva, Switzerland*

***Edited by:***
*Angelo Maravita, University of Milano Bicocca, Italy*

***Reviewed by:***
*Ming Meng, Dartmouth College, USA*
*William Hayward, University of Hong Kong, Hong Kong*

****Correspondence:***
*Olaf Blanke, Laboratory of Cognitive Neuroscience, Ecole Polytechnique Fédérale de Lausanne (EPFL), Station 19, 1015 Lausanne, Switzerland.*
*e-mail: olaf.blanke@epfl.ch*

Scientific investigations on the nature of the self have so far focused on high-level mechanisms. Recent evidence, however, suggests that low-level bottom-up mechanisms of multi-sensory integration play a fundamental role in encoding specific components of bodily self-consciousness, such as self-location and first-person perspective (Blanke and Metzinger, 2009). Self-location and first-person perspective are abnormal in neurological patients suffering from out-of-body experiences (Blanke et al., 2004), and can be manipulated experimentally in healthy subjects by imposing multi-sensory conflicts (Lenggenhager et al., 2009). Activity of the temporo-parietal junction (TPJ) reflects experimentally induced changes in self-location and first-person perspective (Ionta et al., 2011), and dysfunctions in TPJ are causally associated with out-of-body experiences (Blanke et al., 2002). We argue that TPJ is one of the key areas for multi-sensory integration of bodily self-consciousness, that its levels of activity reflect the experience of the conscious "I" as embodied and localized within bodily space, and that these mechanisms can be systematically investigated using state of the art technologies such as robotics, virtual reality, and non-invasive neuroimaging.

**Keywords: self consciousness, body, multi-sensory integration, neuroscience robotics**

## BODILY SELF

Some of the most important brain systems of humans are dedicated to the maintenance of the balance between the self and the external environment, by processing and integrating many different bodily sensory inputs (visual, auditory, vestibular, somatosensory, motor, visceral, etc.), and providing an online representation of the body in the world (Damasio, 1999; Gallagher, 2005; Jeannerod, 2006; Blanke and Metzinger, 2009). In this view, the body representation in the brain is a complex crossroad where multisensory information is compounded in order to build the basis for bodily self-consciousness (Haggard et al., 2003; Jeannerod, 2007; Metzinger, 2008). Many behavioral studies over the last two decades have used techniques imposing multi-sensory conflict as a means to manipulate some components of self-consciousness. For example, the "rubber hand illusion" paradigm showed that by manipulating local aspects of body perception, it is possible to induce an illusory sense of ownership of a fake hand (e.g., Botvinick and Cohen, 1998; Pavani et al., 2000; Ehrsson et al., 2004; Tsakiris and Haggard, 2005; Tsakiris et al., 2007; Aimola Davies et al., 2010). In particular, if participants observe a rubber hand being stroked synchronously with their own (hidden) hand, they tend to report self-attribution of the rubber hand, as if it was their own hand. This illusory self-attribution is often accompanied by a "proprioceptive drift" toward the location of the rubber hand. Specifically, participants report a change in where they feel their real hand to be located (review in Tsakiris, 2010). Similarly, if a participant holds one palm against that of someone else and simultaneously strokes the dorsal side of both her/his own and the other's index finger, an illusory feeling of numbness for the other person's finger can be perceived: the so-called "numbness" illusion (Dieguez et al., 2009). Furthermore, it has recently been shown that illusory self-attribution is not limited to the hands, but extends to other body parts including the face (Sforza et al., 2010). For example, the experience of having one's own face touched whilst simultaneously (the spatial and temporal sense) seeing the same action applied to the face of another, elicits the so-called "enfacement" illusion: that is an illusory sense of face ownership is induced and the other's facial features are incorporated into the participant's face (Sforza et al., 2010). All of these findings on illusory self-attribution support the idea that low-level multi-sensory processes can influence bodily self-consciousness. However, the self and bodily self-consciousness is globally associated with the body, rather than with multiple different body parts (Lenggenhager et al., 2007; Metzinger, 2008; Blanke and Metzinger, 2009). Recent behavioral studies showed that, beyond local aspects of body perception and self-attribution (rubber hand illusion, numbness illusion, face illusion), multisensory conflicts can also be used to manipulate more global aspects of body perception (Ehrsson, 2007; Lenggenhager et al., 2007, 2009; Petkova and Ehrsson, 2008; Aspell et al., 2009, 2010). These studies showed that it is possible to investigate more global aspects of bodily self-consciousness and described several different components thereof, such as self-location, first-person perspective, and self-identification.

## ABNORMAL BODILY SELF-CONSCIOUSNESS

A central aspect of global bodily self-consciousness is the sense of where the self is perceived to be located in space, or "self-location." This apparently obvious link between the self and the body can be altered and experienced as being non-body centered. Patients suffering from out-of-body experiences (OBEs) of neurological origin experience themselves as located outside their own bodily boundaries (abnormal self-location), and report looking at their real body from an elevated perspective in extrapersonal space (abnormal first-person perspective; Irwin, 1985; Blanke et al., 2004; Blanke and Mohr, 2005; De Ridder et al., 2007). Investigations into the neural correlates of OBEs provide insights on the multi-sensory nature of self-consciousness (Irwin, 1985; Brugger et al., 1997; Blanke et al., 2002, 2004; Brugger, 2002; Blanke and Mohr, 2005). Clinical studies showed that OBEs are linked to dysfunctions of the temporo-parietal junction (TPJ; Blanke et al., 2004; Blanke and Mohr, 2005), but also frontal, and parietal cortices (Lopez et al., 2010; Heydrich et al., 2011). Furthermore, electrical stimulation of the TPJ induces OBE-like experiences (Penfield, 1955; Blanke et al., 2002; De Ridder et al., 2007), and the TPJ is activated during mental imagery of "disembodied" self-location (Arzy et al., 2006; Blanke et al., 2010). Based on these findings an association between TPJ dysfunction and OBEs has been proposed (Blanke et al., 2002, 2004; Maillard et al., 2004; Blanke and Mohr, 2005; Brandt et al., 2005; De Ridder et al., 2007; see also Ionta et al., 2011). The TPJ is an excellent candidate for integrating multi-sensory bodily information (and self-consciousness), because it is involved in many self-related processes, such as first-person perspective (Ruby and Decety, 2001; Vogeley and Fink, 2003; Vogeley et al., 2004), self/other discrimination (Farrer et al., 2003; Frith, 2005), theory-of-mind (ToM; review in Frith and Frith, 2003), and self-regulation (Heatherton, 2011). Accordingly, a selective impairment in self-other tasks, such as understanding others' beliefs, has been reported in patients with lesions of the TPJ (Samson et al., 2005). Together with other brain regions, TPJ has also been considered as part of a brain network involved in ToM, that is the ability to understand others' intentions, beliefs, and desires (review in Frith and Frith, 2003). In particular, the right TPJ is believed to play a crucial role in the attribution of mental states (e.g., "she wants to be a teacher"), and both left and right TPJ are recruited when participants are asked to imagine the other's mind (Saxe and Wexler, 2005). Furthermore, activity of the left TPJ seems to be selective for verbal descriptions of another person's beliefs, while the right TPJ seems to respond more selectively to non-verbal stimuli (Saxe and Kanwisher, 2003). In addition, the TPJ also plays a central role in processing vestibular information, with a right hemispheric predominance for otolithic inputs and a left hemispheric predominance for inputs from semicircular canals (see Lopez et al., 2008 for review). In monkeys, neurons in the TPJ discharge during vestibular stimulation, during tactile stimulation of face and trunk, and when a stimulus is in close proximity to the body (Grusser et al., 1990; Duhamel et al., 1998; Bremmer et al., 2002). It is likely that bi- and tri-modal neurons in the TPJ encode the multi-sensory matching of vestibular, visual, and tactile information for the full-body, similar to visuo-tactile bimodal neurons in the premotor and intraparietal sulcus that are anchored to body parts, including the hand (Iriki et al., 1996; Graziano et al., 2000; Maravita and Iriki, 2004).

Jointly, the reviewed data on the role of the TPJ in self-location and first-person perspective, as well as processes related to self-other distinction and ToM, reveal that cognitive and multi-sensory perceptual aspects of the self recruit at least partly overlapping neural substrates. More work is necessary to investigate how both crucial aspects of the self (conscious-perceptual, cognitive, as well as conceptual mechanisms of the self) interrelate behaviorally and neurally at the TPJ and beyond (Blanke and Metzinger, 2009).

## FULL-BODY ILLUSIONS AND SELF-CONSCIOUSNESS

The nature of abnormal self-location and self-identification during OBEs provides a unique opportunity to investigate self-consciousness, but generalization of results is rendered difficult by several methodological issues (e.g. sample size, lesion homogeneity, different etiologies, and/or phenomenology, and generalization to the normal brain). In order to better control manipulations of self-consciousness with standardized and repeatable experimental protocols, several studies have recently induced OBE-like illusions in large samples of healthy participants by presenting ambiguous multi-sensory information. In particular, self-location, first-person perspective, and self-identification have been experimentally manipulated in healthy subjects using visuo-tactile conflicts (e.g., Ehrsson, 2007; Lenggenhager et al., 2007, 2009; Petkova and Ehrsson, 2008; Aspell et al., 2009).

Pioneering studies by Lenggenhager et al. (2007) and Ehrsson (2007) induced changes in self-location and self-identification using congruent and incongruent visuo-tactile multi-sensory inputs. Their general approach was adapted and extended from the original procedure of the rubber hand illusion (review in Tsakiris, 2010), with a particular emphasis on the synchrony between visual and tactile information. In the setup used by Lenggenhager et al. (2007), participants viewed their own back through a head-mounted display (HMD) connected to a video-camera positioned behind their body. In this way they could see their back from a visuo-spatial third-person point of view, as if it was a virtual body. Their own back was then touched with a wooden stick (tactile stimulation) and the HMD showed the movement either with or without a delay (synchronous/asynchronous visual stimulation). Thus, the touch (tactile experience) perceived by participants was either synchronous or asynchronous with respect to that viewed on the visually presented body. The congruence between the visual and the tactile stimulation determined changes in bodily self-consciousness. In particular, subjective reports indicated that when the visual and tactile stimulation were synchronous, stronger self-identification with the virtual body and stronger illusory touch were experienced (Lenggenhager et al., 2007). Furthermore, behavioral measurements of self-location were acquired by displacing the participants (blindfolded) from the position where they were standing during the visuo-tactile stimulation, and asking them to return to the initial position. Importantly, the indicated positions shifted away from participants' actual starting location and toward that of the virtual body (Lenggenhager et al., 2007) *only after synchronous stimulation*. Based on these findings, the authors defined the complex of changes in bodily self-consciousness including self-identification and illusory touch, as well as the self-location change toward the virtual body, as a "full-body illusion."

Ehrsson (2007) used a slightly different setup. Similar to the previous study, participants sat on a chair and wore an HMD connected to two cameras positioned behind their back, affording them with a third-person perspective in extrapersonal space. Dissimilar to the previous study, though, the site of tactile stimulation in this study was the chest, and another stick (identical to the one used for the tactile stimulation) was moved up and down in front of the cameras. The seen and the felt movement were again either synchronous or asynchronous. After 2 min of visuo-tactile multi-sensory stimulation, participants completed a questionnaire. Results indicated that only after the synchronous stroking did participants report the experience of "sitting behind their back" and "looking at themselves from this location." Control questions did not show differences in responses across synchronous and asynchronous conditions. Furthermore, physiological measurements (skin conductance response) were higher during a threat toward the virtual body after the synchronous stroking with respect to the asynchronous stroking condition (Ehrsson, 2007).

In both experiments participants looked at their own body from an external perspective, and only after synchronous stroking did they report stronger self-identification with the virtual body (Ehrsson, 2007; Lenggenhager et al., 2007), and changes in self-location biased toward the position of the virtual body (Lenggenhager et al., 2007). The direct comparison between these two approaches (back vs. chest-stroking, standing vs. sitting position, presence vs. absence of view of the contact between the stick and virtual body, etc.) has recently been provided (Lenggenhager et al., 2009). In this study, the participants' body position was held constant whilst the experimenters measured three different components of bodily self-consciousness: self-location, self-identification, and first-person visuo-spatial perspective. To that end, participants were placed in a prone position and wore an HMD connected to a camera such that they could see their body from above. In one condition participants received the tactile stimulation on their chest and saw a moving stick in front of the camera (with the virtual body in the background). In another condition they felt the stroking on their back and saw the virtual body being touched by the same stick. In both conditions the visual and the tactile stimulations were either synchronous or asynchronous. Participants completed the usual questionnaires on self-identification. Furthermore, self-location was measured by asking participants to imagine dropping a ball from their "felt" location, and estimating the amount of time required by the ball to "hit the ground." The response times (RTs) of this "mental ball dropping" (MBD) were recorded. Lenggenhager et al. (2009) showed that during the back-stroking, self-identification and illusory touch (as indicated by the questionnaires) were stronger after synchronous than asynchronous visuo-tactile stimulation. During the chest-stroking, self-identification and illusory touch were weaker during the synchronous than the asynchronous visuo-tactile stimulation. Results of the MBD indicated that RTs were shorter in the synchronous back-stroking than in the comparable chest-stroking condition, suggesting that the felt "height" was affected. Specifically, RT analysis suggested that felt "height" was lower for the back-stroking condition, thus further suggesting that self-location was biased more toward the virtual body (below) during the back-stroking, and more toward the camera (above) during the chest-stroking.

These data corroborated pioneering self-observations by G. M. Stratton who described his own experiences in an similar experimental setup. This classical setup allowed him to induce changes in how he saw and felt his body. He reported changes in the visual first-person perspective and self-location, when walking with a portable device made of mirrors aligned in such a way that the walker (Stratton himself) could see a projection of his body below and in front of him (Stratton, 1899; see also Blanke et al., 2008). The setup projected an online image of his body in his anterior peripersonal space while he was walking in the countryside of California. He reported progressively increasing changes in self-location and self-identification over the time of exposure, further associated with the feelings of "being out-of-body" (Stratton, 1899). Similarly, a comparable spatial conflict between the visual information relative to the moving body and the multi-sensory cues from the real body can be elicited by asking the participants to wear an HMD onto which their body filmed from an elevated perspective is projected, so that they could see their body while walking in the room (Mizumoto and Ishikawa, 2005). Using this setup participants report to experience the self as located at the position of the visual perspective and simultaneously at the location of the visually presented body (Mizumoto and Ishikawa, 2005). Somewhat comparably in the experimental setup used by Lenggenhager et al. (2009) and Ehrsson (2007) participants saw their own body being stroked synchronously or asynchronously. This induced changes in self-identification and self-location that were further modulated by the synchrony between visual and tactile stimulation. On that basis it has been proposed that self-location and self-identification are strongly influenced by the location of the seen touch, and that embodied self-location and the first-person perspective can be transformed to a disembodied or outside-body self-location and third-person perspective as a function of how and where the visuo-tactile stimulation occurs (Lenggenhager et al., 2009). We argue that experimental designs based on visuo-tactile multi-sensory disintegration might lead to alterations of the first-person perspective, and that this could be further facilitated by a more extended use of virtual reality (Tarr and Warren, 2002; Sanchez-Vives and Slater, 2005; Riva, 2007; Slater et al., 2010), and perhaps through repeated and prolonged exposure to such artificial bodily signals (Stratton, 1899).

The work on perturbation of visual field – prisms adaptation (PA) – provided important insights into visuo-spatial processing that may be related to the reviewed experiments (Striemer and Danckert, 2010). According to the classic PA procedure developed by Richard Held and colleagues, participants are asked to repeatedly perform goal-directed movements while wearing prismatic goggles (Held and Freedman, 1963; Redding and Wallace, 1997). Prismatic goggles allow researchers to induce variable optical deviations between the seen and the real target position. Thus, the goal-directed reaching or pointing movements are shifted in the direction of the visual deviation. These adaptations progressively increase with practice and – when the prismatic goggles are removed – this adaptation generally leads to an error in the opposite direction (Held and Freedman, 1963). The PA procedure affects the everyday correlation between motor signals and sensory feedback. The reviewed visuo-tactile procedures using video and virtual reality techniques in order to manipulate bodily self-consciousness, share several similarities with such adaptations induced by prisms, and affect the everyday correlation between tactile, visual, and vestibular signals. More systematic work is needed to evaluate whether adaptations as those during prism studies also occur during visuo-tactile stroking (this is for example suggested by changes in self-location) and whether comparable post-effects exist.

## NEUROSCIENCE ROBOTICS AND THE NEURAL BASES OF SELF-LOCATION AND FIRST-PERSON PERSPECTIVE

The different setups that investigated self-location and first-person perspective by using video projections and visuo-tactile conflicts, showed that it is possible to manipulate some sub-components of bodily self-consciousness (Ehrsson, 2007; Lenggenhager et al., 2007, 2009). However, the temporal and spatial correspondence between the visual and the tactile stimulation in these setups was always applied by the experimenter. As such, precise and repeatable manipulations free from any possible experimenter bias were illusive to achieve. It was therefore necessary to develop more reliable methodological approaches and to precisely monitor and control what the participants feel and see. Moreover, even though it has been shown that self-location and first-person perspective could be experimentally studied, the neural underpinnings have not been investigated, probably due to the difficulty of applying the visuo-tactile multi-sensory conflict in a well-controlled and repeatable manner during brain imaging data acquisition. Robotic systems and virtual reality are the ideal tools to realize such standardized stimulation, and can therefore improve the control in such experimental studies (Blanke and Gassert, 2009). The rapid evolution of computer- and virtual-reality technology over the past decades has provided researchers with novel tools to explore different modalities of human perception and cognitive function. This has allowed researchers to revisit long-known phenomena and sensory illusions in behavioral studies with well-controlled and repeatable stimuli that can easily be manipulated in order to introduce multi-sensory conflicts. These conditions can further be manipulated to explore how humans integrate information from different sensory modalities and how they react to perceptual conflicts (Ellis et al., 1999; Ernst and Banks, 2002; Bertelson et al., 2003; Ernst and Bulthoff, 2004). Such environments have found increasing applications in clinics, e.g., for phobia treatment and neurorehabilitation (Jang et al., 2002; Holden, 2005).

In order to expand the variety of sensory modalities and include haptic perception, researchers performed studies in mixed environments, combining virtual reality with real objects. For example Carlin et al. (1997) used tactile stimulation and virtual reality to treat arachnophobia. More recently, robotic systems, in the form of haptic displays, have been added to such environments, taking advantage of their unique ability to precisely apply tactile stimuli – both temporally and spatially – or render variable dynamic environments for physical interaction under computer control (Wolpert and Flanagan, 2010). Combined with virtual reality, such systems offer the potential to systematically investigate haptic perception and sensorimotor control with the ability to precisely control and modulate factors such as intensity, location, type, and congruency of stimuli. Flanagan and Wing (1997) used a servo-controlled linear actuator to investigate if the central nervous system (CNS) uses internal models to adjust grip force when stabilizing hand-held loads during arm movements. Ernst and Banks used a haptic interface and virtual reality to measure the variance in visual and haptic percepts, and to explore how these percepts are optimally integrated based on their reliability (Ernst and Banks, 2002). While the previous developments have provided greater control over experimental conditions with reduced variability in the presentation of stimuli, they have so far been limited to behavioral studies, and the associated neural correlates and mechanisms remained unexplored. More recent advances combining virtual reality and/or robotics with non-invasive neuroimaging have therefore opened a whole new range of technology and neuroscience-driven avenues to investigate sensory processing and multi-sensory integration (Gassert et al., 2008a,b; Blanke and Gassert, 2009; Annett and Bischof, 2010; Dueñas et al., 2011). The first functional studies with robotic interfaces were carried out over a decade ago with positron emission tomography (PET; Shadmehr and Holcomb, 1997; Krebs et al., 1998), and took advantage of the fact that PET is not susceptible to electromagnetic interference from conventional robotic systems. However, PET requires injection of radioactive tracers, has low temporal resolution (in the order of a minute for oxygen-based studies), and low spatial resolution, making it difficult to differentiate between activation in functionally different areas. The rapid spread and evolution of functional magnetic resonance imaging (fMRI) over the past years, providing whole brain coverage with high spatial and good temporal resolution, have made this imaging method attractive for neuroscience investigations.

The MR environment precludes the use of conventional robotic devices with fMRI, both for safety and compatibility reasons. However, despite these constraints, a study using fMRI, MR-compatible robotics, and visuo-tactile multi-sensory conflict has recently investigated the neural mechanisms of self-location and first-person perspective (Ionta et al., 2011). A robotic device built from MR-compatible materials, sensors, and actuators was embedded in the MR-scanner bed. Participants lay on an ergonomic mattress divided into two parts, holding a robotic stimulator in the center, between the two mattresses. Based on Lenggenhager et al. (2009), the robotic device moved a tactile stimulator along a linear guide located below the back of the subject, driven by an ultrasonic motor over a rack and pinion gear. A tactile stimulation sphere was attached at the output over a flexible spring blade. This ensured a constant pressure on the participants' back and allowed the tactile stimulus to be presented according to a precisely repeatable movement profile. While feeling the tactile stimulation on the back, participants watched videos through MR-compatible video goggles placed in front of their eyes. The videos showed the back of a human body in a prone position, filmed from an elevated perspective, being stroked (visual stroking) synchronously or asynchronously with respect to the tactile stroking performed by the robotic device on the participants' back. In a control condition the human body was hidden, and participants could see only the rod moving up and down in an empty room. By virtue of this computer-controlled robotic device, the spatial and temporal aspects of the visuo-tactile stimulation were precisely controlled during the fMRI sessions within and across participants. After the visuo-tactile stimulation, self-location was estimated using the MBD task (Lenggenhager et al., 2009). Furthermore, participants completed the questionnaire on self-identification (Ehrsson, 2007; Lenggenhager et al., 2007, 2009) adapted from the original one used for the rubber hand illusion (Botvinick and Cohen, 1998). Confirming pilot testing it was found that some participants felt as if they were looking up at the virtual body (concordant with their real orientation) whilst others felt as if they were looking down on their virtual body (even if they were facing upward). This finding indicated that two different directions of first-person perspective were adopted by participants: those forming the "up-group" had the impression of looking upward, those in the "down-group" of looking downward at the virtual body. Extending the difference in experienced direction of the first-person perspective between both groups (as indicated by subjective reports), behavioral results showed that RTs in the MBD task were significantly different between the synchronous and the asynchronous visuo-tactile stroking only when a human body was observed (not during control conditions). Most importantly, the direction of this effect was different for the two groups: in the up-group self-location was higher during the synchronous condition (longer RTs in the MBD) with respect to the asynchronous condition; in the down-group self-location was lower during the synchronous condition (shorter RTs in the MBD) with respect to the asynchronous condition. Moreover, independently of the synchrony of stroking, participants from the up-group had faster RTs than those in the down-group, suggesting further differences in self-location between the two groups: subjects in the up-group experienced lower height than those in the down-group. These findings indicated that self-location as measured by the MBD was altered in opposite directions in the two groups, depending on the experienced direction of the first-person perspective (subjective reports). fMRI results showed that the activation patterns in TPJ reflected changes in self-location and first-person perspective. In particular, in both groups the magnitude of the BOLD response was lower in conditions with a higher self-location as quantified by the MBD task, and conditions with a lower self-location were associated with a higher BOLD response. Thus, TPJ activity reflected synchrony-related changes in self-location with respect to the position or level of self-location, and further depended differently on the direction of the first-person perspective. Comparable changes in self-location and the direction of the first-person perspective reported by patients with OBEs due to TPJ damage (Ionta et al., 2011) also concur with these behavioral and fMRI data, independent of any potential attention modulation as shown by the effects of stroking synchrony and especially the effect of first-person perspective. OBE patients classically report an elevated perspective that is distant from the body and down-looking (comparable with participants from the down-group). The results obtained in healthy participants are therefore compatible with clinical data in neurological patients with OBEs (Blanke et al., 2004; De Ridder et al., 2007) and reveal that the temporo-parietal cortex, especially in the right hemisphere, encodes these aspects of bodily self-consciousness.

Finally, the (right lateralized) TPJ has been also considered as part of the brain network involved in visuo-spatial attention (review in Corbetta and Shulman, 2011). Interestingly, improvements in visuo-spatial neglect, a pathological condition that typically affects the egocentric spatial relationship with visuo-spatial perspective or extrapersonal space (Karnath, 1994; Farrell and Robertson, 2000; Vogeley and Fink, 2003), are reported following exposure to prisms (Rode et al., 2006), and further extend to other sensory modalities such as touch (Maravita et al., 2003) and hearing (Jacquin-Courtois et al., 2010). Based on these findings, it has been proposed that PA may influence the activity of some visuomotor structures included in the dorsal visual stream and supposed to further mediate both motor and attentional processes (Corbetta and Shulman, 2002; Milner and Goodale, 2006). This interpretation supports the influence that PA might have on perceptual processes based on the interaction between areas of the dorsal and ventral visual stream – superior temporal gyrus (STG) and inferior parietal lobe (Sarri et al., 2006). Indeed visuo-spatial neglect has been linked to TPJ, including STG (Karnath et al., 2001; Halligan et al., 2003), and neglect patients with lesioned TPJ show deficits also in stimulus-driven reorienting attention (Rengachary et al., 2011). Yet, the exact role of TPJ in spatial attention is still controversial based on data in healthy subjects showing that stimulus-driven attentional processes recruit in addition to the right TPJ (Shulman et al., 2010) also insula, and inferior and medial frontal gyri (Corbetta and Shulman, 2002). Conversely, it has been reported that TPJ activity may also decrease in visual attention tasks (Shulman et al., 1997; Gusnard and Raichle, 2001). On the other hand, the activation of TPJ during egocentric visuo-spatial perspective changes (Maguire et al., 1998; Vallar et al., 1999; Ruby and Decety, 2001), and during social perception tasks (Narumoto et al., 2001; Winston et al., 2002) is consistent with clinical and experimental data in self-related processes (Blanke et al., 2004; Blanke and Arzy, 2005; Blanke et al., 2005). In summary there seems to be a functional overlap in the TPJ between processes related to attention and bodily self-consciousness associated with bilateral recruitment in experimental work in healthy subjects and right lateralized TPJ recruitment in patient studies.

## CONCLUSION AND PERSPECTIVES

Here we have reviewed behavioral (Ehrsson, 2007; Lenggenhager et al., 2007, 2009) brain imaging (Arzy et al., 2006; Ionta et al., 2011) and clinical evidence (Brugger et al., 1997; Blanke et al., 2004; Blanke and Mohr, 2005; De Ridder et al., 2007) about three aspects of bodily self-consciousness: self-location, first-person perspective, and self-identification. Clinical findings showed that these three components are dissociable, suggesting that they rely on different neural bases. Behavioral studies showed that such dissociation can be experimentally induced also in healthy subjects via the imposition of multi-sensory conflicts. Brain imaging evidence showed that, as a multi-sensory body-related integration area, the TPJ is involved in all these three aspects of bodily self-consciousness. However, it is worth noting that also other areas including the precuneus (Northoff and Bermpohl, 2004), as well as the prefrontal (Gusnard et al., 2001; Ionta et al., 2010), somatosensory (Ruby and Decety, 2001), and the vestibular cortex (Lopez et al., 2008) are expected to contribute to bodily self-consciousness. Furthermore, recent studies showed the importance of proprioception (Palluel et al., 2011), acoustic information (Aspell et al., 2010), and pain perception (Hansel et al., 2011). Based on the reviewed findings, we conclude that multi-sensory integration is a key brain mechanism for self-consciousness. We suggest that future work should not only investigate mechanisms of visuo-tactile integration, but also their interaction with vestibular, proprioceptive, and cognitive motor signals (i.e., Kannape et al., 2010). We finally suggest that only by using a multi-disciplinary approach combining behavioral and cognitive neuroscience, engineering, and virtual reality with neuroimaging, will it become possible to unravel the detailed mechanisms of bodily self-consciousness and other aspects of self-consciousness.

## REFERENCES

Aimola Davies, A. M., White, R. C., Thew, G., Aimola, N. M., and Davies, M. (2010). Visual capture of action, experience of ownership, and the illusion of self-touch: a new rubber hand paradigm. *Perception* 39, 830–838.

Annett, M. K., and Bischof, W. F. (2010). Investigating the application of virtual reality systems to psychology and cognitive neuroscience research. *Presence Teleop. Virt.* 19, 131–141.

Arzy, S., Thut, G., Mohr, C., Michel, C. M., and Blanke, O. (2006). Neural basis of embodiment: distinct contributions of temporoparietal junction and extrastriate body area. *J. Neurosci.* 26, 8074–8081.

Aspell, J. E., Lavanchy, T., Lenggenhager, B., and Blanke, O. (2010). Seeing the body modulates audiotactile integration. *Eur. J. Neurosci.* 31, 1868–1873.

Aspell, J. E., Lenggenhager, B., and Blanke, O. (2009). Keeping in touch with one's self: multisensory mechanisms of self-consciousness. *PLoS ONE* 4, e6488. doi:10.1371/journal.pone.0006488

Bertelson, P., Vroomen, J., and De Gelder, B. (2003). Visual recalibration of auditory speech identification: a McGurk aftereffect. *Psychol. Sci.* 14, 592–597.

Blanke, O., and Arzy, S. (2005). The out-of-body experience: disturbed self-processing at the temporoparietal junction. *Neuroscientist.* 11, 16–24.

Blanke, O., and Gassert, R. (2009). Total control in virtual reality and robotics. *Front. Neurosci.* 3, 110–111.

Blanke, O., Ionta, S., Fornari, E., Mohr, C., and Maeder, P. (2010). Mental imagery for full and upper human bodies: common right hemisphere activations and distinct extrastriate activations. *Brain Topogr.* 23, 321–332.

Blanke, O., Landis, T., Spinelli, L., and Seeck, M. (2004). Out-of-body experience and autoscopy of neurological origin. *Brain* 127, 243–258.

Blanke, O., and Metzinger, T. (2009). Full-body illusions and minimal phenomenal selfhood. *Trends Cogn. Sci. (Regul. Ed.)* 13, 7–13.

Blanke, O., Metzinger, T., and Lenggenhager, B. (2008). Olaf Blanke et al's response to Kaspar Meyer's E-letter. *Sci. E Lett.*

Blanke, O., and Mohr, C. (2005). Out-of-body experience, heautoscopy, and autoscopic hallucination of neurological origin implications for neurocognitive mechanisms of corporeal awareness and self-consciousness. *Brain Res. Brain Res. Rev.* 50, 184–199.

Blanke, O., Mohr, C., Michel, C. M., Pascual-Leone, A., Brugger, P., Seeck, M., Landis, T., and Thut, G. (2005). Linking out-of-body experience and self processing to mental own-body imagery at the temporoparietal junction. *J. Neurosci.* 25, 550–557.

Blanke, O., Ortigue, S., Landis, T., and Seeck, M. (2002). Stimulating illusory own-body perceptions. *Nature* 419, 269–270.

Botvinick, M., and Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature* 391, 756.

Brandt, C., Brechtelsbauer, D., Bien, C. G., and Reiners, K. (2005). Out-of-body experience as possible seizure symptom in a patient with a right parietal lesion. *Nervenarzt* 76, 1259, 1261–1252.

Bremmer, F., Klam, F., Duhamel, J. R., Ben Hamed, S., and Graf, W. (2002). Visual-vestibular interactive responses in the macaque ventral intraparietal area (VIP). *Eur. J. Neurosci.* 16, 1569–1586.

Brugger, P. (2002). Reflective mirrors: perspective-taking in autoscopic phenomena. *Cogn. Neuropsychiatry* 7, 179–194.

Brugger, P., Regrad, M., and Landis, T. (1997). Illusory reduplication of own's own body: phenomenology and classification of autoscopic phenomena. *Cogn. Neuropsychiatry* 19–38.

Carlin, A. S., Hoffman, H. G., and Weghorst, S. (1997). Virtual reality and tactile augmentation in the treatment of spider phobia: a case report. *Behav. Res. Ther.* 35, 153–158.

Corbetta, M., and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215.

Corbetta, M., and Shulman, G. L. (2011). Spatial neglect and attention networks. *Annu. Rev. Neurosci.* 34, 569–599.

Damasio, A. (1999). *The Feeling of What Happens.* London: Random House.

De Ridder, D., Van Laere, K., Dupont, P., Menovsky, T., and Van De Heyning, P. (2007). Visualizing out-of-body experience in the brain. *N. Engl. J. Med.* 357, 1829–1833.

Dieguez, S., Mercier, M. R., Newby, N., and Blanke, O. (2009). Feeling numbness for someone else's finger. *Curr. Biol.* 19, R1108–R1109.

Dueñas, J., Chapuis, C., Pfeiffer, C., Martuzzi, R., Ionta, S., Blanke, O., and Gassert, R. (2011). Neuroscience robotics to investigate multisensory integration and bodily awareness. *Proc. IEEE Eng. Med. Biol. Conf.* 8348–8352.

Duhamel, J. R., Colby, C. L., and Goldberg, M. E. (1998). Ventral intraparietal area of the macaque: congruent visual and somatic response properties. *J. Neurophysiol.* 79, 126–136.

Ehrsson, H. H. (2007). The experimental induction of out-of-body experiences. *Science* 317, 1048.

Ehrsson, H. H., Spence, C., and Passingham, R. E. (2004). That's my hand! Activity in premotor cortex reflects feeling of ownership of a limb. *Science* 305, 875–877.

Ellis, R. R., Flanagan, J. R., and Lederman, S. J. (1999). The influence of visual illusions on grasp position. *Exp. Brain Res.* 125, 109–114.

Ernst, M. O., and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 429–433.

Ernst, M. O., and Bulthoff, H. H. (2004). Merging the senses into a robust percept. *Trends Cogn. Sci. (Regul. Ed.)* 8, 162–169.

Farrell, M. J., and Robertson, I. H. (2000). The automatic updating of egocentric spatial relationships and its impairment due to right posterior cortical lesions. *Neuropsychologia* 38, 585–595.

Farrer, C., Franck, N., Georgieff, N., Frith, C. D., Decety, J., and Jeannerod, M. (2003). Modulating the experience of agency: a positron emission tomography study. *Neuroimage* 18, 324–333.

Flanagan, J. R., and Wing, A. M. (1997). The role of internal models in motion planning and control: evidence from grip force adjustments during movements of hand-held loads. *J. Neurosci.* 17, 1519–1528.

Frith, C. (2005). The self in action: lessons from delusions of control. *Conscious. Cogn.* 14, 752–770.

Frith, U., and Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 459–473.

Gallagher, S. (2005). *How the Body Shapes the Mind.* Oxford, NY: Clarendon Press.

Gassert, R., Burdet, E., and Chinzei, K. (2008a). MRI-compatible robotics. *IEEE Eng. Med. Biol. Mag.* 27, 12–14.

Gassert, R., Burdet, E., and Chinzei, K. (2008b). Opportunities and challenges in MR-compatible robotics: reviewing the history, mechatronic components, and future directions of this technology. *EEE Eng. Med. Biol. Mag.* 27, 15–22.

Gassert, R., Yamamoto, A., Chapuis, D., Dovat, L., Bleuler, H., and Burdet, E. (2006). Actuation methods for applications in MR environments. *Concepts Magn. Reson. Part B Magn. Reson. Eng.* 29B, 191–209.

Graziano, M. S., Cooke, D. F., and Taylor, C. S. (2000). Coding the location of the arm by sight. *Science* 290, 1782–1786.

Grusser, O. J., Pause, M., and Schreiter, U. (1990). Localization and responses of neurones in the parieto-insular vestibular cortex of awake monkeys (*Macaca fascicularis*). *J. Physiol. (Lond.)* 430, 537–557.

Gusnard, D. A., Akbudak, E., Shulman, G. L., and Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 4259–4264.

Gusnard, D. A., and Raichle, M. E. (2001). Searching for a baseline: functional imaging and the resting human brain. *Nat. Rev. Neurosci.* 2, 685–694.

Haggard, P., Taylor-Clarke, M., and Kennett, S. (2003). Tactile perception, cortical representation and the bodily self. *Curr. Biol.* 13, R170–R173.

Halligan, P. W., Fink, G. R., Marshall, J. C., and Vallar, G. (2003). Spatial cognition: evidence from visual neglect. *Trends Cogn. Sci. (Regul. Ed.)* 7, 125–133.

Hansel, A., Lenggenhager, B., von Kanel, R., Curatolo, M., and Blanke, O. (2011). Seeing and identifying with a virtual body decreases pain perception. *Eur. J. Pain.* 15, 874–879.

Heatherton, T. F. (2011). Neuroscience of self and self-regulation. *Annu. Rev. Psychol.* 62, 363–390.

Held, R., and Freedman, S. J. (1963). Plasticity in human sensorimotor control. *Science* 142, 455–462.

Heydrich, L., Lopez, C., Seeck, M., and Blanke, O. (2011). Partial and full own-body illusions of epileptic origin in a child with right temporoparietal epilepsy. *Epilepsy Behav.* 20, 583–586.

Holden, M. K. (2005). Virtual environments for motor rehabilitation: review. *Cyberpsychol. Behav.* 8, 187–211; discussion 212–189.

Ionta, S., Ferretti, A., Merla, A., Tartaro, A., and Romani, G. L. (2010). Step-by-step: the effects of physical practice on the neural correlates of locomotion imagery revealed by fMRI. *Hum. Brain Mapp.* 31, 694–702.

Ionta, S., Heydrich, L., Lenggenhager, B., Mouthon, M., Fornari, E., Chapuis, D., Gassert, R., and Blanke, O. (2011). Multisensory mechanisms in temporo-parietal cortex support self-location and first-person perspective. *Neuron* 70, 363–374.

Iriki, A., Tanaka, M., and Iwamura, Y. (1996). Coding of modified body schema during tool use by macaque postcentral neurones. *Neuroreport* 7, 2325–2330.

Irwin, H. J. (1985). *Flight of Mind: A Psychological Study of the Out-of-Body Experience.* Metuchen, NJ: Scarecrow Press.

Jacquin-Courtois, S., Rode, G., Pavani, F., O'Shea, J., Giard, M. H., Boisson, D., and Rossetti, Y. (2010). Effect of prism adaptation on left dichotic listening deficit in neglect patients: glasses to hear better? *Brain* 133(Pt 3), 895–908.

Jang, D. P., Ku, J. H., Choi, Y. H., Wiederhold, B. K., Nam, S. W., Kim, I. Y., and Kim, S. I. (2002). The development of virtual reality therapy (VRT) system for the treatment of acrophobia and therapeutic case. *IEEE Trans. Inf. Technol. Biomed.* 6, 213–217.

Jeannerod, M. (2006). *Motor Cognition: What Actions Tell the Self.* Oxford, NY: Oxford University Press.

Jeannerod, M. (2007). Being oneself. *J. Physiol. Paris* 101, 161–168.

Kannape, O. A., Schwabe, L., Tadi, T., and Blanke, O. (2010). The limits of agency in walking humans. *Neuropsychologia* 48, 1628–1636.

Karnath, H. O. (1994). Subjective body orientation in neglect and the interactive contribution of neck muscle proprioception and vestibular stimulation. *Brain* 117(Pt 5), 1001–1012.

Karnath, H. O., Ferber, S., and Himmelbach, M. (2001). Spatial awareness is a function of the temporal not the posterior parietal lobe. *Nature* 411, 950–953.

Krebs, H. I., Brashers-Krug, T., Rauch, S. L., Savage, C. R., Hogan, N., Rubin, R. H., Fischman, A. J., and Alpert, N. M. (1998). Robot-aided functional imaging: application to a motor learning study. *Hum. Brain Mapp.* 6, 59–72.

Lenggenhager, B., Mouthon, M., and Blanke, O. (2009). Spatial aspects of bodily self-consciousness. *Conscious. Cogn.* 18, 110–117.

Lenggenhager, B., Tadi, T., Metzinger, T., and Blanke, O. (2007). Video ergo sum: manipulating bodily self-consciousness. *Science* 317, 1096–1099.

Lopez, C., Halje, P., and Blanke, O. (2008). Body ownership and embodiment: vestibular and multisensory mechanisms. *Neurophysiol. Clin.* 38, 149–161.

Lopez, C., Heydrich, L., Seeck, M., and Blanke, O. (2010). Abnormal self-location and vestibular vertigo in a patient with right frontal lobe epilepsy. *Epilepsy Behav.* 17, 289–292.

Maguire, E. A., Burgess, N., Donnett, J. G., Frackowiak, R. S., Frith, C. D., and O'Keefe, J. (1998). Knowing where and getting there: a human navigation network. *Science* 280, 921–924.

Maillard, L., Vignal, J. P., Anxionnat, R., and Taillandiervespignani, L. (2004). Semiologic value of ictal autoscopy. *Epilepsia* 45, 391–394.

Maravita, A., and Iriki, A. (2004). Tools for the body (schema). *Trends Cogn. Sci. (Regul. Ed.)* 8, 79–86.

Maravita, A., McNeil, J., Malhotra, P., Greenwood, R., Husain, M., and Driver, J. (2003). Prism adaptation can improve contralesional tactile perception in neglect. *Neurology* 60, 1829–1831.

Metzinger, T. (2008). Empirical perspectives from the self-model theory of subjectivity: a brief summary with examples. *Prog. Brain Res.* 168, 215–245.

Milner, A. D., and Goodale, M. A. (2006). *The Visual Brain in Action.* New York: Oxford University Press.

Mizumoto, M., and Ishikawa, M. (2005). Immunity to error through misidentification and the bodily illusion experiment. *J. Conscious. Stud.* 12, 3–19.

Narumoto, J., Okada, T., Sadato, N., Fukui, K., and Yonekura, Y. (2001). Attention to emotion modulates fMRI activity in human right superior temporal sulcus. *Brain Res. Cogn. Brain Res.* 12, 225–231.

Northoff, G., and Bermpohl, F. (2004). Cortical midline structures and the self. *Trends Cogn. Sci. (Regul. Ed.)* 8, 102–107.

Palluel, E., Aspell, J. E., and Blanke, O. (2011). Leg muscle vibration modulates bodily self-consciousness: integration of proprioceptive, visual, and tactile signals. *J. Neurophysiol.* 105, 2239–2247.

Pavani, F., Spence, C., and Driver, J. (2000). Visual capture of touch: out-of-the-body experiences with rubber gloves. *Psychol. Sci.* 11, 353–359.

Penfield, W. (1955). The twenty-ninth Maudsley lecture: the role of the temporal cortex in certain psychical phenomena. *J. Ment. Sci.* 101, 451–465.

Petkova, V. I., and Ehrsson, H. H. (2008). If I were you: perceptual illusion of body swapping. *PLoS ONE* 3, e3832. doi:10.1371/journal.pone.0003832

Redding, G. M., and Wallace, B. (1997). Prism adaptation during target pointing from visible and nonvisible starting locations. *J. Mot. Behav.* 29, 119–130.

Rengachary, J., He, B. J., Shulman, G. L., and Corbetta, M. (2011). A behavioral analysis of spatial neglect and its recovery after stroke. *Front. Hum. Neurosci.* 5:29. doi:10.3389/fnhum.2011.00029

Riva, G. (2007). Virtual reality and telepresence. *Science* 318, 1240–1242.

Rode, G., Klos, T., Courtois-Jacquin, S., Rossetti, Y., and Pisella, L. (2006). Neglect and prism adaptation: a new therapeutic tool for spatial cognition disorders. *Restor. Neurol. Neurosci.* 24, 347–356.

Ruby, P., and Decety, J. (2001). Effect of subjective perspective taking during simulation of action: a PET investigation of agency. *Nat. Neurosci.* 4, 546–550.

Samson, D., Apperly, I. A., Kathirgamanathan, U., and Humphreys, G. W. (2005). Seeing it my way: a case of a selective deficit in inhibiting self-perspective. *Brain* 128(Pt 5), 1102–1111.

Sanchez-Vives, M. V., and Slater, M. (2005). From presence to consciousness through virtual reality. *Nat. Rev. Neurosci.* 6, 332–339.

Sarri, M., Kalra, L., Greenwood, R., and Driver, J. (2006). Prism adaptation changes perceptual awareness for chimeric visual objects but not for chimeric faces in spatial neglect after right-hemisphere stroke. *Neurocase* 12, 127–135.

Saxe, R., and Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind." *Neuroimage* 19, 1835–1842.

Saxe, R., and Wexler, A. (2005). Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychologia* 43, 1391–1399.

Sforza, A., Bufalari, I., Haggard, P., and Aglioti, S. M. (2010). My face in yours: Visuo-tactile facial stimulation influences sense of identity. *Soc. Neurosci.* 5, 148–162.

Shadmehr, R., and Holcomb, H. H. (1997). Neural correlates of motor memory consolidation. *Science* 277, 821–825.

Shulman, G. L., Corbetta, M., Buckner, R. L., Raichle, M. E., Fiez, J. A., Miezin, F. M., and Petersen, S. E. (1997). Top-down modulation of early sensory cortex. *Cereb. Cortex* 7, 193–206.

Shulman, G. L., Pope, D. L., Astafiev, S. V., McAvoy, M. P., Snyder, A. Z., and Corbetta, M. (2010). Right hemisphere dominance during spatial selective attention and target detection occurs outside the dorsal frontoparietal network. *J. Neurosci.* 30, 3640–3651.

Slater, M., Spanlang, B., Sanchez-Vives, M. V., and Blanke, O. (2010). First person experience of body transfer in virtual reality. *PLoS ONE* 5, e10564. doi:10.1371/journal.pone.0010564

Stratton, G. M. (1899). The spatial harmony of touch and sight. *Mind* 8, 492–550.

Striemer, C. L., and Danckert, J. A. (2010). Through a prism darkly: re-evaluating prisms and neglect. *Trends Cogn. Sci. (Regul. Ed.)* 14, 308–316.

Tarr, M. J., and Warren, W. H. (2002). Virtual reality in behavioral neuroscience and beyond. *Nat. Neurosci.* 5(Suppl.), 1089–1092.

Tsakiris, M. (2010). My body in the brain: a neurocognitive model of body-ownership. *Neuropsychologia* 48, 703–712.

Tsakiris, M., and Haggard, P. (2005). The rubber hand illusion revisited: visuotactile integration and self-attribution. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 80–91.

Tsakiris, M., Hesse, M. D., Boy, C., Haggard, P., and Fink, G. R. (2007). Neural signatures of body ownership: a sensory network for bodily self-consciousness. *Cereb. Cortex* 17, 2235–2244.

Vallar, G., Lobel, E., Galati, G., Berthoz, A., Pizzamiglio, L., and Le Bihan, D. (1999). A fronto-parietal system for computing the egocentric spatial frame of reference in humans. *Exp. Brain Res.* 124, 281–286.

Vogeley, K., and Fink, G. R. (2003). Neural correlates of the first-person-perspective. *Trends Cogn. Sci. (Regul. Ed.)* 7, 38–42.

Vogeley, K., May, M., Ritzl, A., Falkai, P., Zilles, K., and Fink, G. R. (2004). Neural correlates of first-person perspective as one constituent of human self-consciousness. *J. Cogn. Neurosci.* 16, 817–827.

Winston, J. S., Strange, B. A., O'Doherty, J., and Dolan, R. J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nat. Neurosci.* 5, 277–283.

Wolpert, D. M., and Flanagan, J. R. (2010). Q&A: Robotics as a tool to understand the brain. *BMC Biol.* 8, 92.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 12 July 2011; paper pending published: 28 September 2011; accepted: 05 December 2011; published online: 23 December 2011.*

*Citation: Ionta S, Gassert R and Blanke O (2011) Multi-sensory and sensorimotor foundation of bodily self-consciousness – an interdisciplinary approach. Front. Psychology 2:383. doi: 10.3389/fpsyg.2011.00383*

*This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.*

frontiers in
PSYCHOLOGY

REVIEW ARTICLE
published: 21 March 2011
doi: 10.3389/fpsyg.2011.00046

# Uncovering multisensory processing through non-invasive brain stimulation

**Nadia Bolognini[1,2]* and Angelo Maravita[1]**

[1] Department of Psychology, University of Milano–Bicocca, Milan, Italy
[2] Neuropsychological Laboratory, Istituto di Recovero e Cura a Carattere Scientifico Istituto Auxologico Italiano, Milan, Italy

***Edited by:***
*Gregor Thut, University of Glasgow, UK*

***Reviewed by:***
*Juha Silvanto, Helsinki University of Technology, Finland*
*Vincenzo Romei, University of Glasgow, UK*

****Correspondence:***
*Nadia Bolognini, Department of Psychology, University of Milano–Bicocca, Building U6, Piazza dell'Ateneo Nuovo 1, 20126 Milan, Italy.*
*e-mail: nadia.bolognini@unimib.it*

Most of current knowledge about the mechanisms of multisensory integration of environmental stimuli by the human brain derives from neuroimaging experiments. However, neuroimaging studies do not always provide conclusive evidence about the causal role of a given area for multisensory interactions, since these techniques can mainly derive correlations between brain activations and behavior. Conversely, techniques of non-invasive brain stimulation (NIBS) represent a unique and powerful approach to inform models of causal relations between specific brain regions and individual cognitive and perceptual functions. Although NIBS has been widely used in cognitive neuroscience, its use in the study of multisensory processing in the human brain appears a quite novel field of research. In this paper, we review and discuss recent studies that have used two techniques of NIBS, namely transcranial magnetic stimulation and transcranial direct current stimulation, for investigating the causal involvement of unisensory and heteromodal cortical areas in multisensory processing, the effects of multisensory cues on cortical excitability in unisensory areas, and the putative functional connections among different cortical areas subserving multisensory interactions. The emerging view is that NIBS is an essential tool available to neuroscientists seeking for causal relationships between a given area or network and multisensory processes. With its already large and fast increasing usage, future work using NIBS in isolation, as well as in conjunction with different neuroimaging techniques, could substantially improve our understanding of multisensory processing in the human brain.

**Keywords: multisensory processing, brain, TMS, tDCS, vision, touch, audition**

## INTRODUCTION

Although multisensory integration has been widely investigated in animals (Stein and Stanford, 2008) and humans (Calvert, 2001; Driver and Noesselt, 2008; Macaluso and Maravita, 2010), only recent research work has started to gain deeper knowledge into the causal involvement of different brain regions, thanks to non-invasive brain stimulation (NIBS) techniques.

In the present paper we will first give an overview on the available techniques of NIBS as valid means for modulating brain activity. Then, we will review research that implied the use of transcranial magnetic stimulation (TMS) to clarify basic aspects of multisensory integration. We will show how single-pulse TMS (sTMS) critically reveals multisensory influence on visual cortical excitability. Second, we will discuss instances in which repetitive TMS (rTMS) was used to induce a temporary interference with brain activity, with the aim of exploring the causal role of heteromodal or unisensory regions in multisensory integration. Finally, we will address the importance of transcranial direct current stimulation (tDCS) for improving multisensory processing.

## NON-INVASIVE BRAIN STIMULATION: BASIC ASPECTS

There are two main techniques of NIBS, namely TMS and tDCS. Both TMS and tDCS appear to be attractive tools for the study of multisensory interactions in the human brain, given that they are non-invasive and safe methods to effectively modulate sensory processing in the cortex. The present paper briefly illustrates the basic technical features of these two techniques (for comprehensive descriptions of NIBS see Pascual-Leone et al., 2000; Priori, 2003; Nitsche et al., 2008; Wassermann et al., 2008; Zaghi et al., 2009; Bolognini and Ro, 2010; Sandrini et al., 2010).

Transcranial magnetic stimulation relies upon the properties of electromagnetic induction; a rapidly changing magnetic field is generated when a high voltage current is passed through a coil. When this coil is held in close proximity to any electrically conducting medium, such as the brain, this time-varying magnetic field induces electrical current that interferes with normal neuronal activity, hence temporarily altering the function of underlying brain areas (see Wassermann et al., 2008, for a comprehensive overview of the physiological mechanisms of TMS). In cognitive studies, TMS has been generally applied with the aim of causing a transient and reversible disruption of cortical activity, and the TMS data have traditionally been interpreted in the "virtual lesion" framework (Pascual-Leone et al., 2000; Walsh and Pascual-Leone, 2003). This terminology was proposed by analogy with neuropsychological and animal lesion studies (Miniussi et al., 2009) and, consistent with this interpretation, TMS has been used in order to define the putative causal role of different cortical areas during the execution of given cognitive and perceptual tasks. However, the interpretation of TMS effects and the establishment of causal links between activity in the targeted brain area and a given behavioral effect is much more complex than suggested by the virtual lesion hypothesis (Miniussi et al., 2009). In fact, the functional effects of TMS can be interpreted only in the context of complex interactions between the characteristics of the stimulation, the anatomical/functional properties of the neural system, and its state of activation. Different mechanisms of action have been proposed by which TMS might interfere with ongoing brain activity and ultimately with behavior: e.g., the suppression of the relevant signal (Harris et al., 2008) or the addition of random neural noise to the ongoing processing (Silvanto and Muggleton, 2008; Ruzzoli et al., 2010; for other relevant studies that have challenged the virtual lesion view, see Siebner et al., 2009; Ziemann, 2010). Adopting a psychophysical approach appears essential to dissociate these two mechanisms and, in turn, to link in a specific way the effects of TMS on behavioral performance and its putative mechanisms of action at the neural level (Miniussi et al., 2009). Another confounding factor in interpreting TMS data is that the relative predominance of either behavioral facilitation or suppression is dependent on the initial activation state of cortical neurons (state-dependency). This state-dependency is critical since the neural impact of any external stimulus represents an interaction with the ongoing brain activity at the time of stimulation (Silvanto and Pascual-Leone, 2008; Silvanto et al., 2008).

The spatial resolution of TMS is highly dependent upon the shape of the stimulating coil, but it can be in the order of a few millimeters (e.g., when using figure-of-eight coil, with circular components of 45 mm), and focal enough to stimulate regions as small as individual fingers representations in the primary motor cortex (Ro et al., 1999).

The coil can be positioned over the brain either functionally, e.g., for motor cortex, searching for the area where the activation of the contralateral hand muscles is induced (Rossini et al., 1994), or by choosing known anatomical landmarks, e.g., the primary somatosensory area of the hand is located at about 3 cm posterior to the motor hotspot for the contralateral hand (Bolognini and Maravita, 2007). With frameless stereotaxic systems, the TMS coil can be navigated more precisely to target specific anatomical sites based on individual subjects' structural (Paus, 1999; Sparing et al., 2010) and functional (Sack et al., 2009) brain images.

The temporal resolution of TMS depends upon the employed stimulation parameters. When sTMS is used, the temporal resolution can be very high and can provide information about brain function in the order of milliseconds, thus allowing the assessment of chronometry of cortical processing (Pascual-Leone et al., 2000; Bolognini and Ro, 2010). If one does not have a temporal hypothesis about when to deliver a single pulse, a different approach would be to use rTMS, which consists of the application of rhythmic trains of multiple TMS pulses. When using rTMS, stimulation frequency seems to be the key parameter determining the direction of the effects. From a physiological point of view, low-frequency rTMS (≤1 Hz) usually results in decreased cortical excitability, whereas at higher frequencies (>1 Hz) cortical excitability is usually increased (Wassermann et al., 2008; Ziemann et al., 2008). Typically, high-frequency rTMS protocols are applied either as a single short train of pulses or several trains with different inter-train intervals, while low-frequency rTMS is typically given as a prolonged continuous stimulation (Bolognini and Ro, 2010; Sandrini et al., 2010). Noteworthy, as mentioned earlier, the effects of rTMS (i.e., facilitation versus inhibition) are not uniquely related to pulse frequency, but they also depend on the initial state of the stimulated brain region (Silvanto and Pascual-Leone, 2008).

There are two main protocols for delivering rTMS. During the "on-line" approach, subjects perform the task and, at stimulus presentation, or with a specific interval preceding or following it, a train of TMS pulses is given to a particular area of the brain. Another popular approach is to stimulate at the site of interest for some minutes before starting a cognitive task (the so-called "off-line" rTMS protocol). Indeed, a crucial feature of rTMS is that it seems capable of changing the activity in a brain area even beyond the duration of the rTMS application itself (Bolognini and Ro, 2010; Sandrini et al., 2010).

A more recent, alternative protocol is "theta burst TMS," in which short bursts of 50-Hz rTMS are repeated at a rate in the theta range (5 Hz). Inhibitory or excitatory effects of this type of stimulation can be obtained using both continuous or intermittent delivery of theta bursts, according to the stimulation parameters (Huang et al., 2005).

Finally, a new emerging paradigm is the use of rhythmic TMS at frequencies mimicking brain rhythms recorded through electroencephalography (EEG) for the study of brain oscillations (Sauseng et al., 2009; Thut and Miniussi, 2009; Miniussi and Thut, 2010; Romei et al., 2010). There is much correlational evidence showing that activity in specific frequency bands is linked to specific cognitive functions. Hence, rhythmic TMS provides the opportunity to entrain brain oscillations. Then, if behavioral modulation is concurrently observed, one can draw causal links between synchronization in functional networks and specific aspects of a task (Miniussi and Thut, 2010). In this context, EEG can be used to study how TMS interacts with rhythmic brain activity, and/or how rhythmic brain stimulation can be used to modify brain functions.

Through TMS is possible not only to modulate the neural activity at the site of stimulation, but also reveal the functional connectivity between different cortical areas by means of paired pulse TMS (Koch and Rothwell, 2009). In this paradigm, two TMS pulses are delivered by two separate coils. A conditioning stimulus is delivered to a brain site, followed by a test stimulus on a different site, on the same or opposite hemisphere. What is measured is the effect of the conditioning stimulus on the response to the test stimulus and, depending on the intensity of the conditioning stimulus and the interstimulus interval, both facilitation and inhibition can be obtained (Ziemann et al., 1996). The effect of the conditioning pulse can change during the execution of a given task, thus providing clues on the causal role of putative intracortical connections over time for that specific task.

The other main method of NIBS, tDCS, is a form of brain polarization that uses a prolonged low-intensity electric current (1–2 mA), delivered to the scalp through two large electrodes (usually 5 cm × 7 cm; Nitsche et al., 2008). tDCS has been used in humans since the distant past, but a reappraisal of this technique with claims of behavioral effects and clinical benefit took place at the turn of this century (for an historical overview, see Priori, 2003). tDCS is now becoming an attractive tool for cognitive neuroscientists, especially in the context of modulating cortical excitability to facilitate skill acquisition, learning, and neural plasticity in the human brain (Wassermann and Grafman, 2005; Nitsche et al., 2008; Bolognini et al., 2009b).

Transcranial direct current stimulation can up- or down-regulate neural activity in the stimulated regions. Increased excitability of the underlying neurons occurs with anodal stimulation, while decreased excitability is seen after cathodal stimulation. With only 13 min of tDCS stimulation, effects on neural excitability outlast the period of stimulation by up to 90 min (Nitsche and Paulus, 2001). tDCS does not stimulate neurons directly, thus increasing their firing, as TMS does. Rather, it most likely targets neuronal signaling by influencing the permeability of ion channels or shifting electrical gradients, therefore modulating the resting membrane threshold (Ardolino et al., 2005). Chemical neurotransmission, either pre- or post-synaptic, may also play a role in tDCS effects (Liebetanz et al., 2002). tDCS has some advantages over TMS, such as its safer and easier application, and the possibility of applying a sham stimulation in a truly double-blind fashion. On the other hand, tDCS has low spatial resolution as compared to TMS, especially when using focal figure-of-eight coil. Yet, computer-based modeling studies of tDCS indicate that maximum current density magnitudes are located beneath the electrodes, at the cortical level (Wagner et al., 2007). Accordingly, despite the very weak current used, DC polarization delivered to specific cortical areas can alter physiological, perceptual, and higher-order cognitive processes in a pretty selective way (Priori, 2003; Wassermann and Grafman, 2005; Nitsche et al., 2008; Bolognini et al., 2009b; Zaghi et al., 2009).

Overall, TMS and tDCS are rapidly becoming essential tools available to neuroscientists for assessing brain functions. As we will illustrate in this review, TMS can be used to determine whether a brain area is causing some multisensory effects or to measure crossmodal changes in cortical excitability, hence nicely complementing other neuroimaging techniques, such as event-related potential (ERPs) and functional magnetic resonance imaging (fMRI). On the other hand, tDCS appears a valuable tool to modulate cortical excitability to facilitate crossmodal interactions, with potential long-lasting modulatory effects on multisensory perception.

## MULTISENSORY INFLUENCES ON VISUAL CORTICAL EXCITABILITY AS REVEALED BY TMS: EVIDENCE FOR CROSSMODAL INTERACTIONS IN PRIMARY VISUAL AREAS

Perception has traditionally been viewed as a modular function with the different sensory modalities operating largely as separate and independent processes. In this view, sensory information is believed to remain isolated by modality within primary sensory areas and the merging of sensory experience is the results of processing occurring within higher-order "heteromodal" associative areas of the brain (Calvert, 2001). However, recent evidence has forced to reconsider this oversimplification. Results from both anatomical and physiological studies suggest that crossmodal interactions occur not only within regions deemed heteromodal, but also within primary sensory areas, i.e., areas traditionally considered to be located very early along the cortical processing stream and receiving direct input from single sensory modalities (Schroeder and Foxe, 2005; Macaluso, 2006; Driver and Noesselt, 2008). This conclusion is strengthened by recent works investigating the crossmodal modulation of visual phosphenes induced by sTMS. The application of sTMS to the occipital visual areas in the human brain can elicit phosphenes, i.e., bright spots of light in specific regions of the visual field (Fernandez et al., 2002; McKeefry et al., 2009). Phosphenes are generated within coextensive regions of the cortex and could be induced by application of TMS to virtually all early visual areas, including the striate cortex (V1), extrastriate areas (V2/V3), and cortico-cortical tracts projecting from V2/V3 back to V1 (Kammer et al., 2005a,b). The features of phosphenes elicited by the stimulation of a given site seem to reflect the perceptual specialization of that area (McKeefry et al., 2009) For instance, while phosphenes elicited by TMS of V1 are stationary, those evoked by TMS of visual motion area V5/MT are often moving (Pascual-Leone et al., 2005).

The TMS output threshold needed to generate phosphenes is believed to provide a direct measure of the excitability of low-level visual cortex (Fernandez et al., 2002; Kammer et al., 2005a). For this reason, one can directly assess crossmodal effects in visual cortex by testing the changes of TMS intensity needed for inducing phosphenes, following a concurrent stimulation in another sensory modality. Using phosphenes, instead of presenting external visual stimuli, may have the advantage of testing crossmodal effects more directly, since phosphenes are generated by direct cortical stimulation, thus bypassing peripheral and subcortical pathways.

Using this approach, it has been shown that a peripheral somatosensory stimulus can modify visual cortex excitability in such a way that phosphenes perception can be now induced using a lower TMS intensity (Ramos-Estebanez et al., 2007). The facilitatory effect of touch over phosphene perception holds a high degree of spatial specificity, since the advantage occurs only for touches delivered exactly at the same spatial location where phosphenes are reported, following the stimulation of the contralateral visual cortex (Bolognini and Maravita, 2007; Ramos-Estebanez et al., 2007). This effect is in line with the crossmodal spatial congruency effects found using fMRI. For instance, Macaluso et al. (2000) demonstrated that a tactile stimulus enhances activity within unimodal visual cortical areas, but only when it is on the same side as the visual target (Macaluso et al., 2000). Moreover, thanks to the optimal temporal resolution of sTMS, the optimal temporal window for the tactile enhancement of phosphenes was found for somatosensory stimuli preceding the occipital sTMS by 60 ms (Ramos-Estebanez et al., 2007), in broadly agreement with ERP evidence for crossmodal tactile modulation of visual responses (Eimer, 2001).

Not only somatosensory, but even auditory inputs can boost visual cortical excitability. Romei et al. (2009) first showed that structured looming sounds, which were of a duration below the psychophysical discrimination threshold, considerably enhance phosphenes perception, as compared to other sound categories (e.g., sounds moving away from the observer or stationary). The onset of the enhancement effect started when TMS followed the auditory stimulus by 80 ms. These findings are of interest since they indicate that cortical excitability in low-level visual areas is rapidly and efficiently boosted by sounds through early, pre-perceptual, and stimulus-selective modulation of neuronal excitability (Romei et al., 2009). Subsequently, it was shown that: (1) the auditory enhancement of phosphenes perception requires a strict spatial correspondence between the two stimuli; (2) the effect of auditory signals occurs only for phosphenes generated in the peripheral visual field (>30°), but not in the central visual field; (3) audio–visual facilitation is present only at subthreshold TMS intensity for phosphenes induction, suggesting that crossmodal interactions depend on the relative physiological salience of the visual information (Bolognini et al., 2010c). Under these conditions, the auditory modulation of phosphenes was maximal when the auditory stimulus preceded the occipital TMS pulse by 40 ms. This timing is in agreement with ERPs recording in humans showing relatively early auditory influences on visual processing (<50 ms from stimulus onset; Giard and Peronnet, 1999; Molholm et al., 2002). Yet, under different experimental conditions, a slightly different latency was found, with an initial auditory modulation of phosphenes starting at 60–80 ms after the sound onset (Romei et al., 2007, 2009). A similar shift in time of audio–visual effects in early visual cortices was also described in a recent ERPs' study by Cappe et al. (2010), who described multisensory interactions occurring at 60–95 ms after audio–visual stimuli. Multisensory effects were shown to be simultaneously originating within a network including primary auditory cortices, primary visual cortices, and posterior superior temporal regions (Cappe et al., 2010).

Overall, the available evidence indicates that auditory inputs can enhance visual cortical excitability over a longer time window, depending on the auditory stimulus selectivity, the perceptual gain, the temporal and spatial features of the combined stimuli. Noteworthy, Romei et al. (2007) also found that sTMS over the occipital pole produced opposing behavioral effects during a simple reaction time task to visual and auditory stimuli, with TMS slowing down reaction times to visual stimuli, but facilitating reaction times to auditory stimuli (Romei et al., 2009). This evidence indicates the existence of bidirectional influences between the stimulus-evoked auditory and TMS-induced visual cortex activities, suggesting state-dependent effects of TMS in the context of multisensory interactions.

Collectively, the above evidence of crossmodal modulation of phosphenes suggests the presence of early multisensory effects, i.e., somatosensory-driven or auditory-driven sensitivity changes in low-level visual areas. This is compatible with the existence of specific pathways linking specialized areas across sensory modalities. These sensory interactions can be revealed under subthreshold conditions and follow the principles of spatial and temporal congruency. The auditory or tactile input might be transmitted to unisensory visual areas through one of two possible pathways (Macaluso, 2006; Driver and Noesselt, 2008). The first pathway consists of a direct, feedforward projections from primary or associative auditory/somatosensory cortices to early visual areas (Falchier et al., 2002; Rockland and Ojima, 2003; Cappe and Barone, 2005; Schroeder and Foxe, 2005; Cappe et al., 2009). In this view, early crossmodal interactions could originate at the level of primary or secondary visual regions. Alternatively, an indirect pathway may be implicated in which feedforward auditory inputs reach areas of multisensory convergence (e.g., superior temporal polysensory region or posterior parietal cortex (PPC), respectively for auditory and tactile input) and are then transmitted via feedback connections to earlier unisensory visual areas (Jones and Powell, 1970; Meredith and Stein, 1983; Giard and Peronnet, 1999; Calvert, 2001; Macaluso, 2006; Meienbrock et al., 2007; Driver and Noesselt, 2008).

Guided by the available evidence from neurophysiologic studies in animals and brain-imaging studies in humans, the studies of crossmodal modulation of phosphenes appear a valuable approach to understand how crossmodal interactions alter visual neuronal excitability in the human low-level sensory cortices. On the other hand, phosphenes can be reported by subjects, but cannot be directly quantified by observers (Boroojerdi et al., 2002). Therefore the use of adequate psychophysical approaches seems important (e.g., Bolognini and Maravita, 2007; Romei et al., 2009). Hence, future work on crossmodal phosphenes might benefit from the combined used of neuroimaging techniques in order to measures the functional effects of the interactions between external stimulus-evoked activity and TMS-induced visual cortex activity.

## TMS DISRUPTION OF MULTISENSORY PROCESSES

Transcranial magnetic stimulation allows the investigation of causality in the brain–behavior relationship, by temporarily altering the activity of neurons in brain areas that are underneath the magnetic field (Bolognini and Ro, 2010). In this context, an advantage of TMS over other neuroimaging methods is that TMS can be used to demonstrate that a brain region is causally essential for performing a given task. In this view, sTMS or rTMS have been used to establish whether a putative heteromodal area of the human brain is essential for multisensory processing.

### AUDIO–VISUAL INTERACTIONS WITH RESPECT TO EXTRAPERSONAL SPACE REPRESENTATION AND SPEECH PERCEPTION

The representation of the space around us is intrinsically multisensory. For instance, the occurrence of sensory signals from multiple sensory modalities, holding close spatial and temporal proximity, are typically integrated in an optimal way since they likely suggest a common source of stimulation (Stein, 1998). Behavioral studies have demonstrated many audio–visual effects in the extrapersonal space: multisensory cues can often either enhance our ability to detect or localize upcoming events (i.e., multisensory enhancement; Gielen et al., 1983; Hughes et al., 1998; Spence et al., 1998; Stein, 1998; Bolognini et al., 2005) or bias the localization of unisensory events (i.e., ventriloquism; Howard and Templeton, 1966; Hairston et al., 2003; Alais and Burr, 2004). As discussed above, these multisensory spatial effects may be subserved by converging information from sensory-specific cortices into multisensory areas, which in turn would affect early unisensory regions via feedback projections (Macaluso, 2006).

Starting from these considerations, TMS was used to explore the role of the PPC in audio–visual spatial interactions. The PPC is a heteromodal region of sensory convergence that contains both neurons responding to isolated visual and auditory stimuli (Bushara et al., 1999; Bremmer et al., 2001) and multisensory neurons (Andersen, 1997). These latter cells may be the ideal locus for multisensory integration, thus contributing to supramodal cognitive functions, such as spatial orienting and spatial awareness (Andersen, 1997). By applying an off-line rTMS paradigm (i.e., 20 min of 1-Hz rTMS, delivered before the task), we assessed whether the rTMS interference over the right inferior parietal lobule (IPL) could cause a disruption of modality-specific spatial orienting and/or a disturbance of the typical response speed advantage observed with crossmodal targets, as measured by the redundant target effect (Maravita et al., 2008). The results showed that low-frequency rTMS over IPL increased reaction times to spatially lateralized modality-specific visual and auditory stimuli, without affecting the response to bimodal audio–visual stimuli. Crucially, a residual advantage for multisensory stimuli, supported by a neural co-activation mechanism, was retained in spite of the parietal interference (Bolognini et al., 2009a). Control rTMS over V1 impaired only contralateral visual responses, without affecting the response to auditory or audio–visual targets. The modality-specific auditory and visual spatial deficits observed after IPL-rTMS could be due to the selective disruption of neighboring unisensory spatial representations for each of the two modalities within the IPL. Indeed, neuroimaging evidence shows modality-specific activations within multiple subregions of the PPC (Andersen, 1997; Bushara et al., 1999; Downar et al., 2000; Bremmer et al., 2001). On the other hand, the persistence of audio–visual speed advantage after disruption suggests that other structures beyond IPL may have compensated any local interference induced by rTMS (Maravita et al., 2008), supporting a still efficient integration of audio–visual inputs.

Bertini et al. (2010) investigated the neural substrate of audio–visual interactions subserving localization abilities by applying continuous theta burst TMS (cTBS, 40 s of cTBS consisting of bursts of three TMS pulses delivered at 50 Hz, with each train burst repeated every 200 ms) before an auditory localization task with unisensory auditory, and spatially congruent and incongruent audio–visual stimuli. The authors found that the visual enhancement of auditory localization induced by congruent audio–visual stimuli was disrupted by cTBS of the temporo-parietal cortex (TPC), whereas the ventriloquism effect (i.e., the perceptual translocation of the sound toward the visual stimulus during incongruent audio–visual stimulation) was reduced by cTBS of the occipital cortex (OC). Even in this task, cTBS of IPL did not affect audio–visual localization (Bertini et al., 2010). Such findings are in broad agreement with both neuroimaging (Calvert, 2001; Bonath et al., 2007) and neuropsychological evidence (Frassinetti et al., 2005; Leo et al., 2008). Unfortunately, the authors did not control for visual suppression effects during occipital cTBS, thus it remains to be clarified whether the reduction of the ventriloquism effect could be ascribed, partially or totally, to a mere reduction of visual sensitivity.

With respect to multisensory integration of speech-related stimuli, a compelling example of multisensory integration is the McGurk effect, in which an auditory syllable is perceived very differently depending on whether it is accompanied by a visual movie of a speaker pronouncing the same syllable or a different, incongruent syllable (McGurk and MacDonald, 1976). Anatomical and physiological studies in human and non-human primates have suggested that the superior temporal sulcus (STS) is involved in audio–visual integration for both speech and non-speech stimuli (Calvert, 2001; Beauchamp et al., 2004). Across three experiments, the causal role of STS in the occurrence of the McGurk effect was tested by Beauchamp et al. (2010) using fMRI-guided sTMS. The chronometry of STS involvement was also assessed by delivering sTMS at different time intervals, i.e., from 298 ms before to 362 ms after the auditory stimulus onset, with steps of 66 ms. The results showed that disruption of the left STS, but not of a control site located dorsally and posteriorly to the STS, significantly reduced the likelihood of the McGurk effect. Such an effect could not be ascribed to a more general impairment of speech perception, as assessed by using control stimuli. Moreover, the McGurk effect was reduced only when sTMS was delivered within a 200-ms time window, spanning 100 ms before to 100 ms after auditory stimulus presentation (Beauchamp et al., 2010). This time–effect is in line with previous electrophysiological evidence demonstrating strong responses in STS beginning around 100 ms after stimulus presentation (Schroeder and Foxe, 2002; Barraclough et al., 2005; Canolty et al., 2007; Puce et al., 2007).

Hence, studies of audio–visual interactions, despite the methodological differences among them, highlight a causal role of heteromodal regions around the temporal cortex, namely TPC for spatial audio–visual interactions, and STS for audio–visual speech-related processing. Some degree of crossmodal spatial interactions also emerged within primary visual areas, although this would benefit from further investigation (Falchier et al., 2002; Ghazanfar and Schroeder, 2006; Driver and Noesselt, 2008). Instead, IPL appears to process both visual and auditory spatial signals, but their integration seems to take place outside this cortical region, at least for relatively low-level orienting response. In the next future, the multisensory functions of other subregions of the PPC (e.g., superior parietal lobule, SPL) should be tested.

## VISUOTACTILE INTERACTIONS WITH RESPECT TO BODY AND PERIPERSONAL SPACE REPRESENTATION

Our continuous interactions with the external world are subtended by specific brain mechanisms controlling the integration between bodily perceptions and spatially localized visual stimuli. In this respect, near peripersonal, as opposed to far extrapersonal space, is defined as the space near the body where action typically occurs and a close control of bodily stimuli (e.g., a touch) and visual stimuli (e.g., an object contacting our body) must be constantly maintained. Specific neural structures contain bimodal neurons responding to visual stimuli near the body. These bimodal neurons are considered critical to visuotactile integration in the peripersonal space in animals (Rizzolatti et al., 1997; Graziano, 2001) and in humans (Maravita et al., 2003; Ladavas and Farnè, 2004). The critical aspect of this kind of multisensory integration is that it is maintained across bodily movements. The hand, for example, keeps interacting with visual stimuli in its surroundings, even if it continuously moves across different spatial positions, as shown by electrophysiological research (Graziano et al., 1997) and research on neurologically healthy subjects and brain-damaged patients (Maravita et al., 2003). Bolognini and Maravita (2007) showed the causal involvement of the PPC in maintaining the constant mapping of visual and tactile reference frames across hand movements. These authors showed that the enhancement of TMS-induced phosphenes by spatially congruent touches (see previous paragraph for details about this experiment) was abolished if off-line 1-Hz rTMS was applied to the right PPC, in a position putatively targeting the ventral intraparietal cortex (area VIP). This area was shown to be specifically activated during the monitoring of tactile input across hand postures in a previous fMRI study (Lloyd et al., 2003). In particular, before rTMS, a touch on the index finger increased the rate of reported phosphenes both when the hands were uncrossed (e.g., detection of phosphenes in the left hemifield following right occipital TMS was increased by left-hand touches) and crossed (i.e., left-hemifield phosphenes now increased following right-hand touches), showing a spatial realignment of visual and somatosensory frames of reference following hand crossing. Critically, following rTMS to VIP, the rate of phosphenes in a given hemifield (e.g., the left) was enhanced by tactile stimuli delivered to the hand (the left, in this example) controlled by the same brain hemisphere (the right one), regardless of the hand posture. In other words, no remapping of visual and somatosensory representations occurred following changes of hand posture, but now crossmodal facilitation was mediated by an "intrahemispheric" facilitation mechanism (left/right hand always increasing the report of phosphenes in the homologous left/right hemifield).

Evidence for the causal involvement of intraparietal cortex in the proprioceptive remapping of touch into external space was recently further confirmed by Azanon et al. (2010) using sTMS.

As outlined in the previous chapter in relation to audio–visual integration, even for visuotactile interactions an important issue is which mechanism subserves the integration of stimuli belonging to different sensory modalities (Macaluso and Maravita, 2010). A recent study by Chambers et al. (2007) has clarified the causal role of the inferior parietal cortex in the reflexive orienting toward visual or somatosensory events, induced by non-informative visual or tactile cues. On-line rTMS was delivered synchronously with cue onset for 100 ms (pulses at 0, 50, and 100 ms post-cue onset). A disruption of the angular gyrus caused a deficit in the exogenous attention orienting induced by a tactile cue on the detection of both tactile and visual targets, in a task requiring an elevation judgment of visual and tactile events delivered close to the hands. This area showed, therefore, a selective role in biasing attention within and across sensory modalities, following a somatosensory, but not a visual cue. By contrast, rTMS to the supramarginal gyrus reduced reflexive orienting only across, but not within, sensory modalities, showing selective crossmodal effects (Chambers et al., 2007).

The above cited studies underline the importance of parietal areas for multisensory integration. On one side, the intraparietal sulcus seems critical to maintain a constant update of limb posture, thus updating the relative position of extrapersonal visual and somatosensory bodily stimuli for visuotactile effects, which are so typical of the peripersonal space (Maravita et al., 2003). Furthermore, attentional orienting toward unisensory and bimodal visual and tactile events are subtended by discrete subregions of the parietal lobe.

### UNCOVERING CROSSMODAL EFFECTS IN UNISENSORY VISUAL OR SOMATOSENSORY AREAS

Another emerging line of investigation aims at uncovering whether a modality-specific cortical region, commonly associated with the processing of unisensory information from a given sense, can be recruited in a crossmodal manner to process purely unisensory stimuli from a different sensory modality. This line of investigation differs from those described above (see paragraph 3), which assessed the occurrence of multisensory interactions among bimodal stimuli, in unisensory and heteromodal cortices. In humans, sensory deprivation has been widely used as a model to explore the role of experience-dependent crossmodal recruitment of modality-specific areas, as revised elsewhere (e.g., Neville and Bavelier, 2002; Merabet and Pascual-Leone, 2010). Here, we will take into consideration only recent studies showing crossmodal processing in traditionally viewed modality-specific areas of the healthy, not sensory-deprived, human brain. TMS evidence of the contribution of visual imagery in the crossmodal recruitment of visual cortex in tactile processing are not considered here, but are reviewed elsewhere (e.g., Sathian and Zangaladze, 2002).

In this context, Mancini et al. (2010) took advantage of the Müller–Lyer illusion (i.e., arrowheads at the ends of a line may affect its estimated length) to address the causal role of the regions activated by the visual illusion in the crossmodal generation of the haptic, and crossmodal visuo-haptic illusory effects (Mancini et al., 2010). Off-line 1-Hz rTMS was administered to the lateral occipital cortex (LOC), and to the right SPL of both hemispheres. SPL is known to be involved in visual (Corbetta et al., 1995) and crossmodal spatial attention and localization (Bushara et al., 1999), whereas LOC is considered a modality-specific visual area, although LOC might be also implicated in crossmodal object recognition, responding to both familiar and unfamiliar visual and tactile shapes (Amedi et al., 2002; Deshpande et al., 2010). rTMS to left and right LOC impaired both unisensory (visual, haptic) and crossmodal processing of the Müller–Lyer illusion in a similar fashion. Conversely, rTMS to left and right SPC did not affect the illusion in any modality. These results showed that the visual area LOC plays a crossmodal role in tactile perception of both illusory and non-illusory shapes; although being traditional identified as a visual area, LOC is involved in the processing of purely somatosensory stimuli, as well as visuo-haptic stimuli.

A further experiment tested whether disruption of somatosensory cortices could impair the processing of visual stimuli depicting a touch. This situation is being recently put forward by different researchers, as a possible somatosensory analog of the mirror system for movement. While the motor mirror system show activation of motor areas by the vision of moving bodily segments (Rizzolatti and Craighero, 2004), its analog in the somatosensory system would show activation of primary (S1) and/or secondary (S2) somatosensory cortices by the vision of a tactile event (Keysers et al., 2010). This system may contribute to the understanding of sensory events occurring in other people (Gallese, 2007), in a broader circuit of embodied simulation, whereby visual events may be mapped onto our own body. This mechanism could involve the activation of multisensory brain areas and be altered if the PPC are disrupted by TMS (Pasalar et al., 2010). Bolognini et al. (2011b) showed that on-line 13-Hz rTMS delivered over S1 disrupted the performance in a visual discrimination task requiring to judge whether a moving finger was or not touching a hand. rTMS to S1 selectively reduced performance for contralateral stimuli. The effect was specific for the perception of touch, since the discrimination of a finger moving, without touch, was unaffected by S1 stimulation (Bolognini et al., 2011b).

The traditional definition of "purely" unisensory areas has been effectively changed by the above described TMS studies, showing that typical unisensory visual areas may exert somatosensory or multisensory functions (Mancini et al., 2010), whereas early somatosensory areas can process purely visual stimuli depicting touches (Bolognini et al., 2011b). These results add novel insights to the neuroimaging literature, showing the causal involvement of unisensory cortices in crossmodal perception.

## FACILITATING MULTISENSORY INTERACTIONS BY tDCS

Ongoing studies in our laboratory are focusing on the use of tDCS to modulate multisensory processing. For instance, the polarization of the right PPC may enhance spatial orienting across different sensory modalities. Specifically, 15 min of anodal tDCS (2 mA, delivered before the task) of the right PPC can decrease response latency to unimodal visual and auditory stimuli, as well as to bimodal audio–visual stimuli, during a reaction time task; the modulation was specific for stimuli contralateral to the tDCS side, with control stimulation of the OC affecting only contralateral visual processing (Bolognini et al., 2010b). As discussed above, low-frequency rTMS over the right PPC induced opposite effects, with increased latency of responses to unisensory visual and auditory stimuli (Bolognini et al., 2009a).

Collectively the two studies, apart from the different NIBS techniques used, converge in showing the causal involvement of the right PPC in supramodal spatial orienting across different sensory modalities. Moreover, a dissociation of the tDCS effects was found with respect to the crossmodal processing of blue audio–visual stimuli, which are likely not detected by the human collicular pathway (Leh et al., 2010), versus red audio–visual stimuli, which are detected by the superior colliculus and the PPC. The tDCS-induced facilitation was indeed stronger for blue audio–visual stimuli, mostly integrated at a cortical level, whereas responses to red audio–visual stimuli, which likely involve a subcortical level of processing (Maravita et al., 2008), appeared less susceptible to the DC polarization of the PPC.

Based on these findings, and given the putative facilitatory effects of anodal tDCS on learning (Reis et al., 2008), we then explored the effect of coupling a multisensory visual field exploration training with anodal tDCS (2 mA for 30 min, delivered during the practice). We found that anodal tDCS delivered to the right, but not left, PPC is effective in facilitating performance during the audio–visual exploration training, inducing a speeding up of responses to audio–visual stimuli in the early phase of practice. Instead, without tDCS, performance improves more gradually and later during training, only at the end of practice. Additionally, the multisensory training combined with anodal tDCS translated in a greater gain in a variety of tasks assessing visuo-spatial attention and search, as compared to the effect induced by the training given in isolation, not coupled with tDCS (Bolognini et al., 2010a).

Crossmodal illusions are the flip side of sensory coherence, and they illustrate some of the consequences of disrupting the normal relationships among different sensory cues (Stein, 1998). One of the most powerful example of multisensory perception is the "sound-induced flash illusion" (Shams et al., 2000). When a single flash is presented along with two or more beeps, observers often report seeing two or more flashes, the so-called "fission" illusion. A corresponding "fusion" illusion has also been described, where a single beep causes the fusion of a double flash stimulus (Andersen et al., 2004). These multisensory phenomena highlight how sensory-specific perceptual judgments concerning one sense (i.e., vision) can be dramatically affected by their interaction with other senses (i.e., audition). Using this audio–visual illusion, we explored the possibility of modulating multisensory perception by polarizing with tDCS putatively relevant cortical regions, which are likely involved in the generation of the illusion (Watkins et al., 2006, 2007). We found that up- or down-regulating cortical excitability by tDCS can facilitate or reduce audio–visual interactions, depending on the current polarity, the targeted area, and the illusory percept. Specifically, the perceptual "fission" of a single flash, due to multiple beeps, was increased after anodal tDCS of the temporal cortex (i.e., active electrode placed at the level of the superior temporal gyrus), and decreased after anodal tDCS of the OC (i.e., active electrode placed at the level of V1). A reversal of such effects was induced by cathodal tDCS. Conversely, the perceptual "fusion" of multiple flashes due to a single beep was unaffected by tDCS and polarizing the parietal cortex was overall ineffective (Bolognini et al., 2011a). This study shows that tDCS can modulate audio–visual interactions, by non-invasively shifting cortical excitability in occipital or in temporal cortices, and it elucidates the causal association between neural activity in the occipital and temporal areas, and the conscious visual experience brought about by a multisensory stimulation.

The available data concerning multisensory processes shows that tDCS may facilitate multisensory interactions in a variety of experimental conditions. The observed facilitation of behavior depends on different factors, including the site of the stimulated area, the tDCS parameters (duration, intensity), the type of task, and also the stage of learning, when a training paradigm is used.

## CONCLUSION AND FUTURE DIRECTIONS

Current evidence suggests that multisensory processing occurs in higher-order heteromodal areas, as well as the earliest stages of sensory processing, up to primary sensory cortices in the human brain. Exactly how our brain integrates inputs from the different senses remains the subject of intensive investigation. It is becoming increasingly clear that many multisensory phenomena may reflect the causal interplay between interconnected regions of the brain, rather than just the function of an isolated area (Macaluso, 2006; Driver and Noesselt, 2008). The chance to directly inquire whether a given area or network has a causal or merely subsidiary role in shaping multisensory interactions can be assuaged by techniques of NIBS, which appear essential tools to substantially improve our understanding of multisensory processes, extending the knowledge derived from neuroimaging studies. Except for the study of crossmodal plasticity after deafferentation, the study of multisensory processing in the healthy human brain by using TMS or tDCS has began quite recently. TMS approaches allow to transiently interfere with the activity in a given area, in order to assess the relationship between focal cortical activity and multisensory-related behavior and to trace the timing at which activity in a particular cortical region contributes to multisensory interactions. On the other hand, tDCS consists of a subthreshold stimulation that modulates cortical excitability, consequently modulating the amount of multisensory effects. Another interesting approach for future studies is to probe the functional connectivity of different cortical areas with the use of paired pulse TMS. Double-pulse paradigms could be successfully used to study the time course of intra- and inter-hemispheric cortico-cortical pathways subserving multisensory interactions. Moreover, TMS can be combined with other techniques for measuring brain function, such as EEG, PET, fMRI, in order to explore the functional interplay between heteromodal and modality-specific brain areas. There is a clear theoretical advantage in combining different approaches. Combining two different methods, such as TMS and neuroimaging, has the advantage of overcoming the intrinsic limitations of either techniques used in isolation. In particular, with this approach it is possible to supplement the information provided by correlational analysis of neuroimaging, with a technique that can establish a causal link between brain function and behavior (Sack and Linden, 2003; Miniussi and Thut, 2010). On the other hand, a great advantage of functional neuroimaging is the ability to acquire measurements of activity in the entire brain, thus providing a broader picture of the cortical responses to TMS; hence, combining TMS with fMRI or PET allows to investigate interregional interactions and their possible functional consequences for perception and cognition (Sack and Linden, 2003; Ruff et al., 2006). Furthermore, the TMS–EEG integration can offer real-time information on cortical reactivity and connectivity, and on how functional activity links to behavior through the study of TMS-induced modulations (Miniussi and Thut, 2010).

We also need to go beyond the standard "virtual lesion" approach, used so far in TMS studies of multisensory processing. Based on the state-dependent effects of TMS, new paradigms were proposed which might discriminate between functionally distinct neuronal representations in the stimulated area.

Finally, although TMS cannot directly target subcortical structures, a recent study has shown that activity in the thalamus can be modulated by stimulation of parietal cortex, hence opening up new horizons for the studies of cortical–subcortical interactions in multisensory processing (Blankenburg et al., 2008).

Whichever approach is adopted, the use of NIBS in the study of multisensory interactions requires carefully designed psychophysical tasks and the testing of control sites, in order to ensure the selectivity of the neuronal disruption of a given area on unisensory versus multisensory processing. In the next future, the knowledge derived from NIBS studies can undoubtedly extend our comprehension of the multisensory mechanisms operating in the human brain.

## ACKNOWLEDGMENT

The preparation of the present paper was supported by Funding (FAR 2010) from the University of Milano–Bicocca to Angelo Maravita and Nadia Bolognini.

## REFERENCES

Alais, D., and Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.* 14, 257–262.

Amedi, A., Jacobson, G., Hendler, T., Malach, R., and Zohary, E. (2002). Convergence of visual and tactile shape processing in the human lateral occipital complex. *Cereb. Cortex* 12, 1202–1212.

Andersen, R. A. (1997). Multimodal integration for the representation of space in the posterior parietal cortex. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 352, 1421–1428.

Andersen, T. S., Tiippana, K., and Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Brain Res. Cogn. Brain Res.* 21, 301–308.

Ardolino, G., Bossi, B., Barbieri, S., and Priori, A. (2005). Non-synaptic mechanisms underlie the after-effects of cathodal transcutaneous direct current stimulation of the human brain. *J. Physiol.* 568, 653–663.

Azanon, E., Longo, M. R., Soto-Faraco, S., and Haggard, P. (2010). The posterior parietal cortex remaps touch into external space. *Curr. Biol.* 20, 1304–1309.

Barraclough, N. E., Xiao, D., Baker, C. I., Oram, M. W., and Perrett, D. I. (2005). Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. *J. Cogn. Neurosci.* 17, 377–391.

Beauchamp, M. S., Argall, B. D., Bodurka, J., Duyn, J. H., and Martin, A. (2004). Unraveling multisensory integration: patchy organization within human STS multisensory cortex. *Nat. Neurosci.* 7, 1190–1192.

Beauchamp, M. S., Nath, A. R., and Pasalar, S. (2010). fMRI-guided transcranial magnetic stimulation reveals that the superior temporal sulcus is a cortical locus of the McGurk effect. *J. Neurosci.* 30, 2414–2417.

Bertini, C., Leo, F., Avenanti, A., and Ladavas, E. (2010). Independent mechanisms for ventriloquism and multisensory integration as revealed by theta-burst stimulation. *Eur. J. Neurosci.* 31, 1791–1799.

Blankenburg, F., Ruff, C. C., Bestmann, S., Bjoertomt, O., Eshel, N., Josephs, O., Weiskopf, N., and Driver, J. (2008). Interhemispheric effect of parietal TMS on somatosensory response confirmed directly with concurrent TMS-fMRI. *J. Neurosci.* 28, 13202–13208.

Bolognini, N., Frassinetti, F., Serino, A., and Ladavas, E. (2005). "Acoustical vision" of below threshold stimuli: interaction among spatially converging audiovisual inputs. *Exp. Brain Res.* 160, 273–282.

Bolognini, N., Fregni, F., Casati, C., Olgiati, E., and Vallar, G. (2010a). Brain polarization of parietal cortex augments training-induced improvement of visual exploratory and attentional skills. *Brain Res.* 1349, 76–89.

Bolognini, N., Olgiati, E., Rossetti, A., and Maravita, A. (2010b). Enhancing multisensory spatial orienting by brain polarization of the parietal cortex. *Eur. J. Neurosci.* 31, 1800–1806.

Bolognini, N., Senna, I., Maravita, A., Pascual-Leone, A., and Merabet, L. B. (2010c). Auditory enhancement of visual phosphene perception: the effect of temporal and spatial factors and of stimulus intensity. *Neurosci. Lett.* 477, 109–114.

Bolognini, N., and Maravita, A. (2007). Proprioceptive alignment of visual and somatosensory maps in the posterior parietal cortex. *Curr. Biol.* 17, 1890–1895.

Bolognini, N., Miniussi, C., Savazzi, S., Bricolo, E., and Maravita, A. (2009a). TMS modulation of visual and auditory processing in the posterior parietal cortex. *Exp. Brain Res.* 195, 509–517.

Bolognini, N., Pascual-Leone, A., and Fregni, F. (2009b). Using non-invasive brain stimulation to augment motor training-induced plasticity. *J. Neuroeng. Rehabil.* 6, 8.

Bolognini, N., and Ro, T. (2010). Transcranial magnetic stimulation: disrupting neural activity to alter and assess brain function. *J. Neurosci.* 30, 9647–9650.

Bolognini, N., Rossetti, A., Casati, C., Mancini, F., and Vallar, G. (2011a). Neuromodulation of multisensory perception: a tDCS study of the sound-induced flash illusion. *Neuropsychologia* 49, 231–237.

Bolognini, N., Rossetti, R., Maravita, A., and Miniussi, C. (2011b). Seeing touches in the somatosensory cortex: a TMS study of the visual perception of touch. *Hum. Brain Mapp.* doi: 10.1002/hbm.21172. [Epub ahead of print].

Bonath, B., Noesselt, T., Martinez, A., Mishra, J., Schwiecker, K., Heinze, H. J., and Hillyard, S. A. (2007). Neural basis of the ventriloquist illusion. *Curr. Biol.* 17, 1697–1703.

Boroojerdi, B., Meister, I. G., Foltys, H., Sparing, R., Cohen, L. G., and Topper, R. (2002). Visual and motor cortex excitability: a transcranial magnetic stimulation study. *Clin. Neurophysiol.* 113, 1501–1504.

Bremmer, F., Schlack, A., Shah, N. J., Zafiris, O., Kubischik, M., Hoffmann, K., Zilles, K., and Fink, G. R. (2001). Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalencies between humans and monkeys. *Neuron* 29, 287–296.

Bushara, K. O., Weeks, R. A., Ishii, K., Catalan, M. J., Tian, B., Rauschecker, J. P., and Hallett, M. (1999). Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. *Nat. Neurosci.* 2, 759–766.

Calvert, G. A. (2001). Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cereb. Cortex* 11, 1110–1123.

Canolty, R. T., Soltani, M., Dalal, S. S., Edwards, E., Dronkers, N. F., Nagarajan, S. S., Kirsch, H. E., Barbaro, N. M., and Knight, R. T. (2007). Spatiotemporal dynamics of word processing in the human brain. *Front. Neurosci.* 1:1. doi: 10.3389/neuro.01/1.1.014.2007

Cappe, C., and Barone, P. (2005). Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. *Eur. J. Neurosci.* 22, 2886–2902.

Cappe, C., Morel, A., Barone, P., and Rouiller, E. M. (2009). The thalamocortical projection systems in primate: an anatomical support for multisensory and sensorimotor interplay. *Cereb. Cortex* 19, 2025–2037.

Cappe, C., Thut, G., Romei, V., and Murray, M. M. (2010). Auditory-visual multisensory interactions in humans: timing, topography, directionality, and sources. *J. Neurosci.* 30, 12572–12580.

Chambers, C. D., Payne, J. M., and Mattingley, J. B. (2007). Parietal disruption impairs reflexive spatial attention within and between sensory modalities. *Neuropsychologia* 45, 1715–1724.

Corbetta, M., Shulman, G. L., Miezin, F. M., and Petersen, S. E. (1995). Superior parietal cortex activation during spatial attention shifts and visual feature conjunction. *Science* 270, 802–805.

Deshpande, G., Hu, X., Lacey, S., Stilla, R., and Sathian, K. (2010). Object familiarity modulates effective connectivity during haptic shape perception. *Neuroimage* 49, 1991–2000.

Downar, J., Crawley, A. P., Mikulis, D. J., and Davis, K. D. (2000). A multimodal cortical network for the detection of changes in the sensory environment. *Nat. Neurosci.* 3, 277–283.

Driver, J., and Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. *Neuron* 57, 11–23.

Eimer, M. (2001). Crossmodal links in spatial attention between vision, audition, and touch: evidence from event-related brain potentials. *Neuropsychologia* 39, 1292–1303.

Falchier, A., Clavagnier, S., Barone, P., and Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *J. Neurosci.* 22, 5749–5759.

Fernandez, E., Alfaro, A., Tormos, J. M., Climent, R., Martinez, M., Vilanova, H., Walsh, V., and Pascual-Leone, A. (2002). Mapping of the human visual cortex using image-guided transcranial magnetic stimulation. *Brain Res. Brain Res. Protoc.* 10, 115–124.

Frassinetti, F., Bolognini, N., Bottari, D., Bonora, A., and Ladavas, E. (2005). Audiovisual integration in patients with visual deficit. *J. Cogn. Neurosci.* 17, 1442–1452.

Gallese, V. (2007). Embodied simulation: from mirror neuron systems to interpersonal relations. *Novartis Found. Symp.* 278, 3–12; discussion 12–19, 89–96, 216–221.

Ghazanfar, A. A., and Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends Cogn. Sci.* 10, 278–285.

Giard, M. H., and Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J. Cogn. Neurosci.* 11, 473–490.

Gielen, S. C., Schmidt, R. A., and Van den Heuvel, P. J. (1983). On the nature of intersensory facilitation of reaction time. *Percept. Psychophys.* 34, 161–168.

Graziano, M. S. (2001). A system of multimodal areas in the primate brain. *Neuron* 29, 4–6.

Graziano, M. S., Hu, X. T., and Gross, C. G. (1997). Coding the locations of objects in the dark. *Science* 277, 239–241.

Hairston, W. D., Wallace, M. T., Vaughan, J. W., Stein, B. E., Norris, J. L., and Schirillo, J. A. (2003). Visual localization ability influences cross-modal bias. *J. Cogn. Neurosci.* 15, 20–29.

Harris, I. M., Benito, C. T., Ruzzoli, M., and Miniussi, C. (2008). Effects of right parietal transcranial magnetic stimulation on object identification and orientation judgments. *J. Cogn. Neurosci.* 20, 916–926.

Howard, I. P., and Templeton, W. B. (1966). *Human Spatial Orientation*. London: Wiley.

Huang, Y. Z., Edwards, M. J., Rounis, E., Bhatia, K. P., and Rothwell, J. C. (2005). Theta burst stimulation of the human motor cortex. *Neuron* 45, 201–206.

Hughes, H. C., Nelson, M. D., and Aronchick, D. M. (1998). Spatial characteristics of visual-auditory summation in human saccades. *Vision Res.* 38, 3955–3963.

Jones, E. G., and Powell, T. P. (1970). An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. *Brain* 93, 793–820.

Kammer, T., Puls, K., Erb, M., and Grodd, W. (2005a). Transcranial magnetic stimulation in the visual system. II. Characterization of induced phosphenes and scotomas. *Exp. Brain Res.* 160, 129–140.

Kammer, T., Puls, K., Strasburger, H., Hill, N. J., and Wichmann, F. A. (2005b). Transcranial magnetic stimulation in the visual system. I. The psychophysics of visual suppression. *Exp. Brain Res.* 160, 118–128.

Keysers, C., Kaas, J. H., and Gazzola, V. (2010). Somatosensation in social perception. *Nat. Rev. Neurosci.* 11, 417–428.

Koch, G., and Rothwell, J. C. (2009). TMS investigations into the task-dependent functional interplay between human posterior parietal and motor cortex. *Behav. Brain Res.* 202, 147–152.

Ladavas, E., and Farnè, A. (2004). Visuotactile representation of near-the-body space. *J. Physiol. Paris* 98, 161–170.

Leh, S. E., Ptito, A., Schonwiesner, M., Chakravarty, M. M., and Mullen, K. T. (2010). Blindsight mediated by an S-cone-independent collicular pathway: an fMRI study in hemispherectomized subjects. *J. Cogn. Neurosci.* 22, 670–682.

Leo, F., Bolognini, N., Passamonti, C., Stein, B. E., and Ladavas, E. (2008). Cross-modal localization in hemianopia: new insights on multisensory integration. *Brain* 131, 855–865.

Liebetanz, D., Nitsche, M. A., Tergau, F., and Paulus, W. (2002). Pharmacological approach to the mechanisms of transcranial DC-stimulation-induced after-effects of human motor cortex excitability. *Brain* 125, 2238–2247.

Lloyd, D. M., Shore, D. I., Spence, C., and Calvert, G. A. (2003). Multisensory representation of limb position in human premotor cortex. *Nat. Neurosci.* 6, 17–18.

Macaluso, E. (2006). Multisensory processing in sensory-specific cortical areas. *Neuroscientist* 12, 327–338.

Macaluso, E., Frith, C. D., and Driver, J. (2000). Modulation of human visual cortex by crossmodal spatial attention. *Science* 289, 1206–1208.

Macaluso, E., and Maravita, A. (2010). The representation of space near the body through touch and vision. *Neuropsychologia* 48, 782–795.

Mancini, F., Bolognini, N., Bricolo, E., and Vallar, G. (2010). Cross-modal processing in the occipito-temporal cortex: a TMS study of the Muller–Lyer illusion. *J. Cogn. Neurosci.* doi:10.1162/jocn.2010.21561. [Epub ahead of print].

Maravita, A., Bolognini, N., Bricolo, E., Marzi, C. A., and Savazzi, S. (2008). Is audiovisual integration subserved by the superior colliculus in humans? *Neuroreport* 19, 271–275.

Maravita, A., Spence, C., and Driver, J. (2003). Multisensory integration and the body schema: close to hand and within reach. *Curr. Biol.* 13, R531–R539.

McGurk, H., and MacDonald, J. (1976). Hearing lips and seeing voices. *Nature* 264, 746–748.

McKeefry, D. J., Gouws, A., Burton, M. P., and Morland, A. B. (2009). The noninvasive dissection of the human visual cortex: using FMRI and TMS to study the organization of the visual brain. *Neuroscientist* 15, 489–506.

Meienbrock, A., Naumer, M. J., Doehrmann, O., Singer, W., and Muckli, L. (2007). Retinotopic effects during spatial audio-visual integration. *Neuropsychologia* 45, 531–539.

Merabet, L. B., and Pascual-Leone, A. (2010). Neural reorganization following sensory loss: the opportunity of change. *Nat. Rev. Neurosci.* 11, 44–52.

Meredith, M. A., and Stein, B. E. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science* 221, 389–391.

Miniussi, C., Ruzzoli, M., and Walsh, V. (2009). The mechanism of transcranial magnetic stimulation in cognition. *Cortex* 46, 128–130.

Miniussi, C., and Thut, G. (2010). Combining TMS and EEG offers new prospects in cognitive neuroscience. *Brain Topogr.* 22, 249–256.

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., and Foxe, J. J. (2002). Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Brain Res. Cogn. Brain Res.* 14, 115–128.

Neville, H., and Bavelier, D. (2002). Human brain plasticity: evidence from sensory deprivation and altered language experience. *Prog. Brain Res.* 138, 177–188.

Nitsche, M. A., Cohen, L. G., Wassermann, E. M., Priori, A., Lang, N., Antal, A., Paulus, W., Hummel, F., Boggio, P. S., Fregni, F., and Pascual-Leone, A. (2008). Transcranial direct current stimulation: state of the art 2008. *Brain Stimul.* 1, 206–223.

Nitsche, M. A., and Paulus, W. (2001). Sustained excitability elevations induced by transcranial DC motor cortex stimulation in humans. *Neurology* 57, 1899–1901.

Pasalar, S., Ro, T., and Beauchamp, M. S. (2010). TMS of posterior parietal cortex disrupts visual tactile multisensory integration. *Eur. J. Neurosci.* 31, 1783–1790.

Pascual-Leone, A., Amedi, A., Fregni, F., and Merabet, L. B. (2005). The plastic human brain cortex. *Annu. Rev. Neurosci.* 28, 377–401.

Pascual-Leone, A., Walsh, V., and Rothwell, J. (2000). Transcranial magnetic stimulation in cognitive neuroscience – virtual lesion, chronometry, and functional connectivity. *Curr. Opin. Neurobiol.* 10, 232–237.

Paus, T. (1999). Imaging the brain before, during, and after transcranial magnetic stimulation. *Neuropsychologia* 37, 219–224.

Priori, A. (2003). Brain polarization in humans: a reappraisal of an old tool for prolonged non-invasive modulation of brain excitability. *Clin. Neurophysiol.* 114, 589–595.

Puce, A., Epling, J. A., Thompson, J. C., and Carrick, O. K. (2007). Neural responses elicited to face motion and vocalization pairings. *Neuropsychologia* 45, 93–106.

Ramos-Estebanez, C., Merabet, L. B., Machii, K., Fregni, F., Thut, G., Wagner, T. A., Romei, V., Amedi, A., and Pascual-Leone, A. (2007). Visual phosphene perception modulated by subthreshold crossmodal sensory stimulation. *J. Neurosci.* 27, 4178–4181.

Reis, J., Robertson, E., Krakauer, J. W., Rothwell, J., Marshall, L., Gerloff, C., Wassermann, E., Pascual-Leone, A., Hummel, F., Celnik, P. A., Classen, J., Floel, A., Ziemann, U., Paulus, W., Siebner, H. R., Born, J., and Cohen, L. G. (2008). Consensus: "can tDCS and TMS enhance motor learning and memory formation?". *Brain Stimul.* 1, 363–369.

Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192.

Rizzolatti, G., Fadiga, L., Fogassi, L., and Gallese, V. (1997). The space around us. *Science* 277, 190–191.

Ro, T., Cheifet, S., Ingle, H., Shoup, R., and Rafal, R. (1999). Localization of the human frontal eye fields and motor hand area with transcranial magnetic stimulation and magnetic resonance imaging. *Neuropsychologia* 37, 225–231.

Rockland, K. S., and Ojima, H. (2003). Multisensory convergence in calcarine visual areas in macaque monkey. *Int. J. Psychophysiol.* 50, 19–26.

Romei, V., Gross, J., and Thut, G. (2010). On the role of prestimulus alpha rhythms over occipito-parietal areas in visual input regulation: correlation or causation? *J. Neurosci.* 30, 8692–8697.

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

Romei, V., Murray, M. M., Merabet, L. B., and Thut, G. (2007). Occipital transcranial magnetic stimulation has opposing effects on visual and auditory stimulus detection: implications for multisensory interactions. *J. Neurosci.* 27, 11465–11472.

Rossini, P. M., Barker, A. T., Berardelli, A., Caramia, M. D., Caruso, G., Cracco, R. Q., Dimitrijević, M. R., Hallett, M., Katayama, Y., Lücking, C. H., de Noordhout, M., Marsden, C. D., Murray, N. M. F., Rothwell, J. C., Swash, M., and Tomberg, C. (1994). Non-invasive electrical and magnetic stimulation of the brain, spinal cord and roots: basic principles and procedures for routine clinical application. Report of an IFCN committee. *Electroencephalogr. Clin. Neurophysiol.* 91, 79–92.

Ruff, C. C., Blankenburg, F., Bjoertomt, O., Bestmann, S., Freeman, E., Haynes, J. D., Rees, G., Josephs, O., Deichmann, R., and Driver, J. (2006). Concurrent TMS-fMRI and psychophysics reveal frontal influences on human retinotopic visual cortex. *Curr. Biol.* 16, 1479–1488.

Ruzzoli, M., Marzi, C. A., and Miniussi, C. (2010). The neural mechanisms of the effects of transcranial magnetic stimulation on perception. *J. Neurophysiol.* 103, 2982–2989.

Sack, A. T., Cohen Kadosh, R., Schuhmann, T., Moerel, M., Walsh, V., and Goebel, R. (2009). Optimizing functional accuracy of TMS in cognitive studies: a comparison of methods. *J. Cogn. Neurosci.* 21, 207–221.

Sack, A. T., and Linden, D. E. (2003). Combining transcranial magnetic stimulation and functional imaging in cognitive brain research: possibilities and limitations. *Brain Res. Brain Res. Rev.* 43, 41–56.

Sandrini, M., Umilta, C., and Rusconi, E. (2010). The use of transcranial magnetic stimulation in cognitive neuroscience: a new synthesis of methodological issues. *Neurosci. Biobehav. Rev.* 35, 516–536.

Sathian, K., and Zangaladze, A. (2002). Feeling with the mind's eye: contribution of visual cortex to tactile perception. *Behav. Brain Res.* 135, 127–132.

Sauseng, P., Klimesch, W., Heise, K. F., Gruber, W. R., Holz, E., Karim, A. A., Glennon, M., Gerloff, C., Birbaumer, N., and Hummel, F. C. (2009). Brain oscillatory substrates of visual short-term memory capacity. *Curr. Biol.* 19, 1846–1852.

Schroeder, C. E., and Foxe, J. (2005). Multisensory contributions to low-level, 'unisensory' processing. *Curr. Opin. Neurobiol.* 15, 454–458.

Schroeder, C. E., and Foxe, J. J. (2002). The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Brain Res. Cogn. Brain Res.* 14, 187–198.

Shams, L., Kamitani, Y., and Shimojo, S. (2000). Illusions. What you see is what you hear. *Nature* 408, 788.

Siebner, H. R., Hartwigsen, G., Kassuba, T., and Rothwell, J. C. (2009). How does transcranial magnetic stimulation modify neuronal activity in the brain? Implications for studies of cognition. *Cortex* 45, 1035–1042.

Silvanto, J., Muggleton, N., and Walsh, V. (2008). State-dependency in brain stimulation studies of perception and cognition. *Trends Cogn. Sci.* 12, 447–454.

Silvanto, J., and Muggleton, N. G. (2008). New light through old windows: moving beyond the "virtual lesion" approach to transcranial magnetic stimulation. *Neuroimage* 39, 549–552.

Silvanto, J., and Pascual-Leone, A. (2008). State-dependency of transcranial magnetic stimulation. *Brain Topogr.* 21, 1–10.

Sparing, R., Hesse, M. D., and Fink, G. R. (2010). Neuronavigation for transcranial magnetic stimulation (TMS): where we are and where we are going. *Cortex* 46, 118–120.

Spence, C., Nicholls, M. E., Gillespie, N., and Driver, J. (1998). Cross-modal links in exogenous covert spatial orienting between touch, audition, and vision. *Percept. Psychophys.* 60, 544–557.

Stein, B. E. (1998). Neural mechanisms for synthesizing sensory information and producing adaptive behaviors. *Exp. Brain Res.* 123, 124–135.

Stein, B. E., and Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nat. Rev. Neurosci.* 9, 255–266.

Thut, G., and Miniussi, C. (2009). New insights into rhythmic brain activity from TMS-EEG studies. *Trends Cogn. Sci.* 13, 182–189.

Wagner, T., Fregni, F., Fecteau, S., Grodzinsky, A., Zahn, M., and Pascual-Leone, A. (2007). Transcranial direct current stimulation: a computer-based human model study. *Neuroimage* 35, 1113–1124.

Walsh, V., and Pascual-Leone, A. (eds). (2003) *Transcranial Magnetic Stimulation: A Neurochronometrics of Mind.* Cambridge, MA: The Mitt Press.

Wassermann, E. M., Epstein, C. M., Ziemann, U., Walsh, V., Paus, T., and Lisanby, S. H. (2008). *The Oxford Handbook of Transcranial Stimulation.* (New York: Oxford University Press).

Wassermann, E. M., and Grafman, J. (2005). Recharging cognition with DC brain polarization. *Trends Cogn. Sci.* 9, 503–505.

Watkins, S., Shams, L., Josephs, O., and Rees, G. (2007). Activity in human V1 follows multisensory perception. *Neuroimage* 37, 572–578.

Watkins, S., Shams, L., Tanaka, S., Haynes, J. D., and Rees, G. (2006). Sound alters activity in human V1 in association with illusory visual perception. *Neuroimage* 31, 1247–1256.

Zaghi, S., Acar, M., Hultgren, B., Boggio, P. S., and Fregni, F. (2009). Noninvasive brain stimulation with low-intensity electrical currents: putative mechanisms of action for direct and alternating current stimulation. *Neuroscientist.* 16, 285–307.

Ziemann, U. (2010). TMS in cognitive neuroscience: virtual lesion and beyond. *Cortex* 46, 124–127.

Ziemann, U., Paulus, W., Nitsche, M. A., Pascual-Leone, A., Byblow, W. D., Berardelli, A., Siebner, H. R., Classen, J., Cohen, L. G., and Rothwell, J. C. (2008). Consensus: motor cortex plasticity protocols. *Brain Stimul.* 1, 164–182.

Ziemann, U., Rothwell, J. C., and Ridding, M. C. (1996). Interaction between intracortical inhibition and facilitation in human motor cortex. *J. Physiol.* 496, 873–881.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 30 November 2010; accepted: 04 March 2011; published online: 21 March 2011.*

*Citation: Bolognini N and Maravita A (2011) Uncovering multisensory processing through non-invasive brain stimulation. Front. Psychology* ***2****:46. doi: 10.3389/fpsyg.2011.00046*

*This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.*