# 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*

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 mature 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 interneu-

ron populations. Moreover, elements in the SC are reciprocally connected by lateral inhibitory or excitatory synapses with a Mexican hat disposition.

## 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;
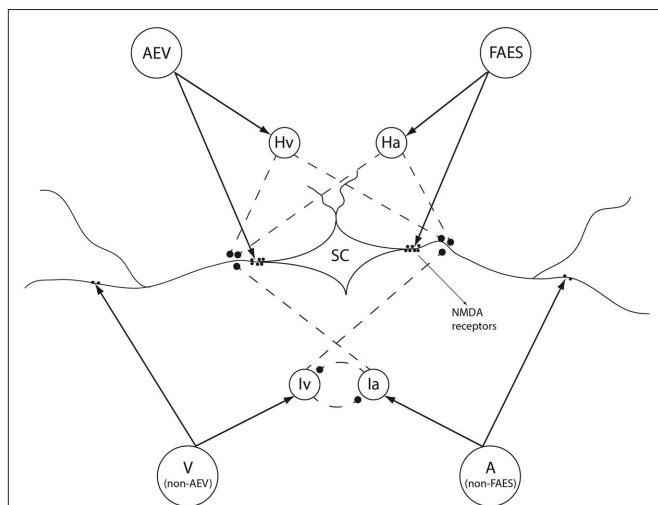


FIGURE 1 | The general structure of the superior colliculus (SC) network. The four projection areas make excitatory synapses with their target SC neurons and with their target interneurons (solid black arrows). The interneurons, by means of inhibitory synapses (dashed black lines), provide two competitive 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.
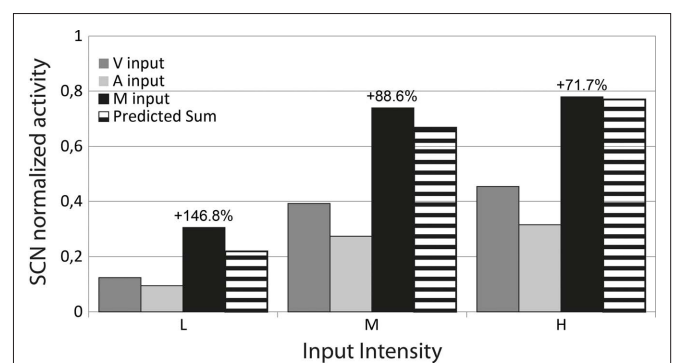


FIGURE 2 | Multisensory Enhancement and Inverse Effectiveness in the model. Activities evoked in the SC neurons in response to Visual (V input, dark gray bars), Auditory (A input, light gray bars) and Cross-modal (M input, black bars) stimuli at different levels of efficacy, placed at the center of the 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
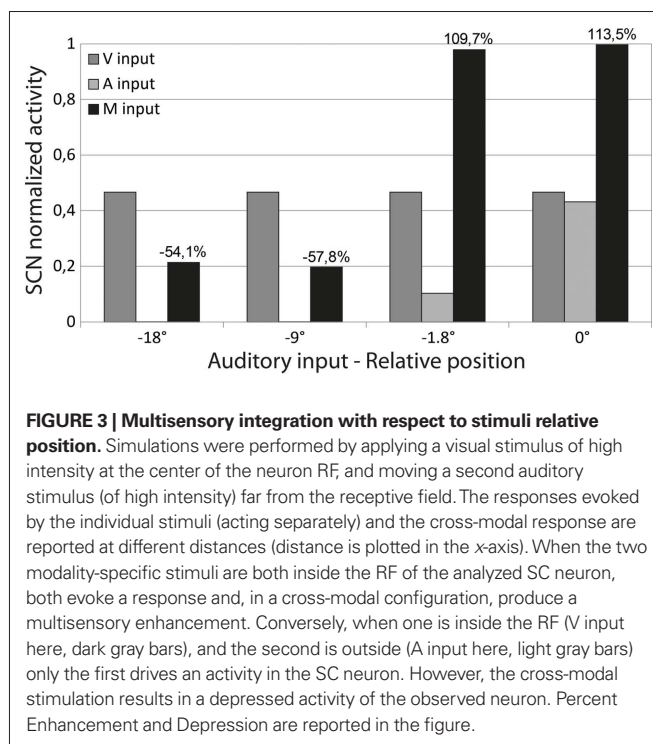


**FIGURE 3 | Multisensory integration with respect to stimuli relative position.** Simulations were performed by applying a visual stimulus of high intensity at the center of the neuron RF, and moving a second auditory stimulus (of high intensity) far from the receptive field. The responses evoked by 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.

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).

### 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
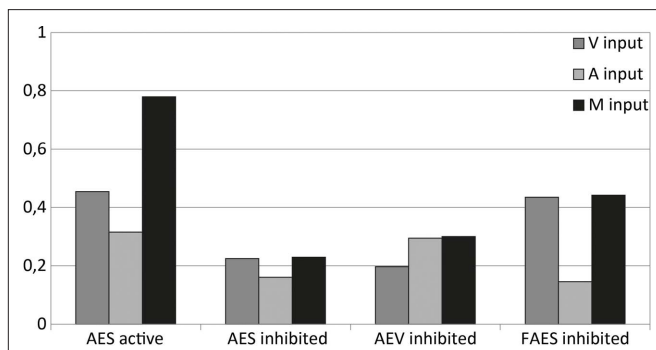
**FIGURE 4 | Behavior of the network as function of AES cortex.** These figures compare the activity of SC neurons in response to different inputs with AES active or inhibited, fully (AES inhibited) or only partially (AEV inhibited, FAES inhibited). In all simulations, the activity was assessed by stimulating 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.
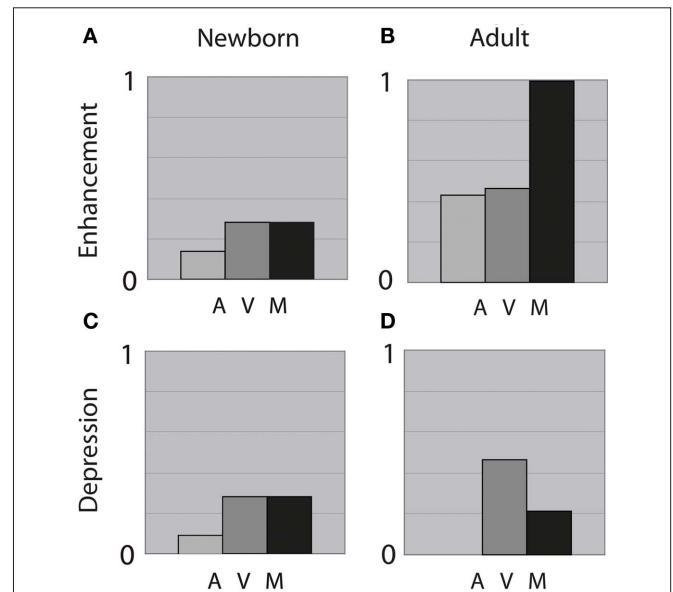
**FIGURE 5 | Integrative capabilities in the neonate and in the adult cats.** **(A,B)** Shows the responses obtained using two stimuli (one auditory and the other visual) of high intensity placed at the center of the RF. **(C,D)** Shows the responses evoked by a visual stimulus at the center of the RF, 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 ($p < 0.0001$).
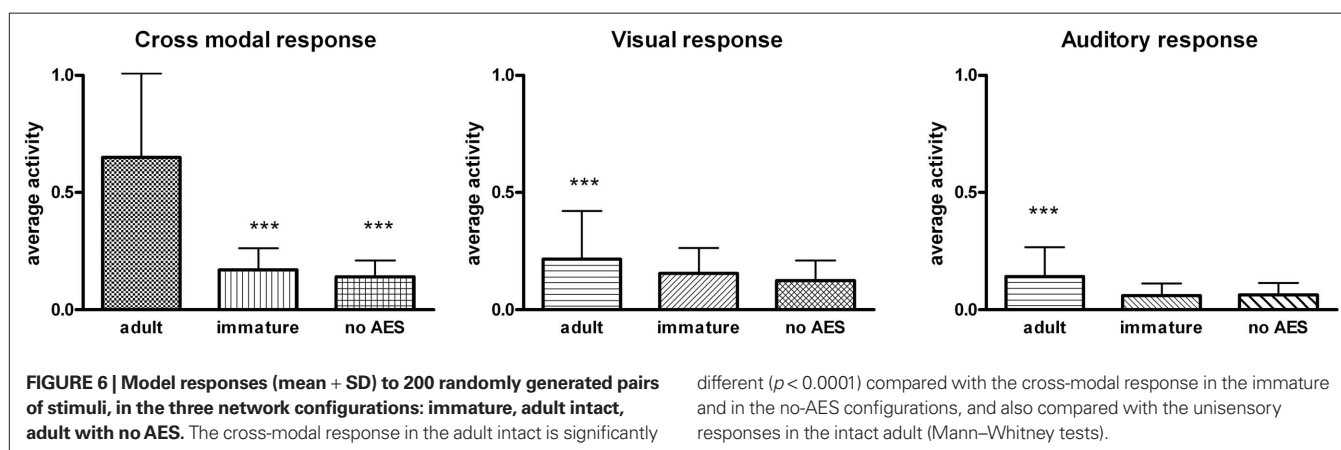
### DISCUSSION

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 mature 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.
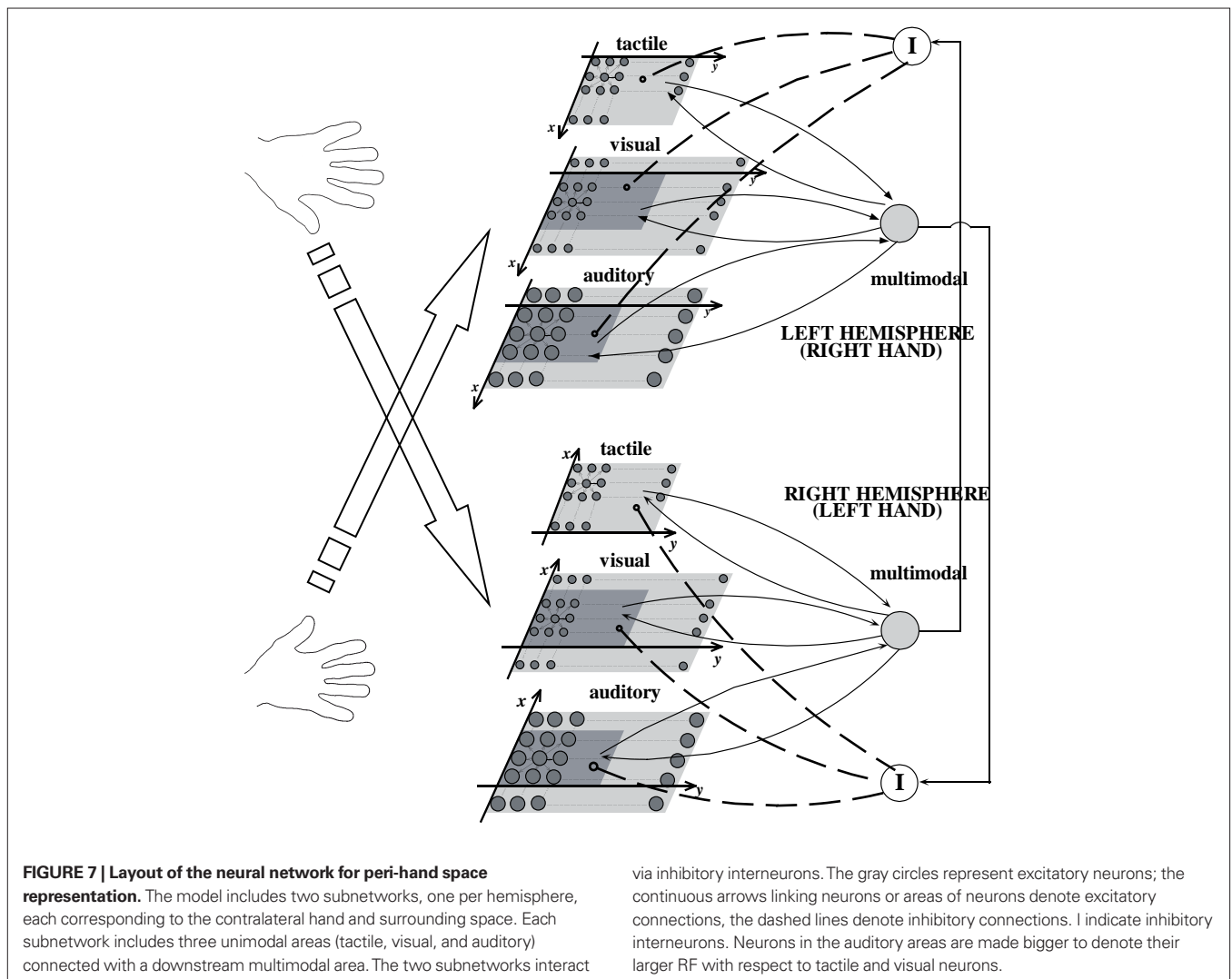
## MODEL DESCRIPTION

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 $10\,cm \times 20\,cm$, roughly representing the surface of the hand. Both the visual and auditory areas cover a space of $15\,cm \times 100\,cm$, 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).

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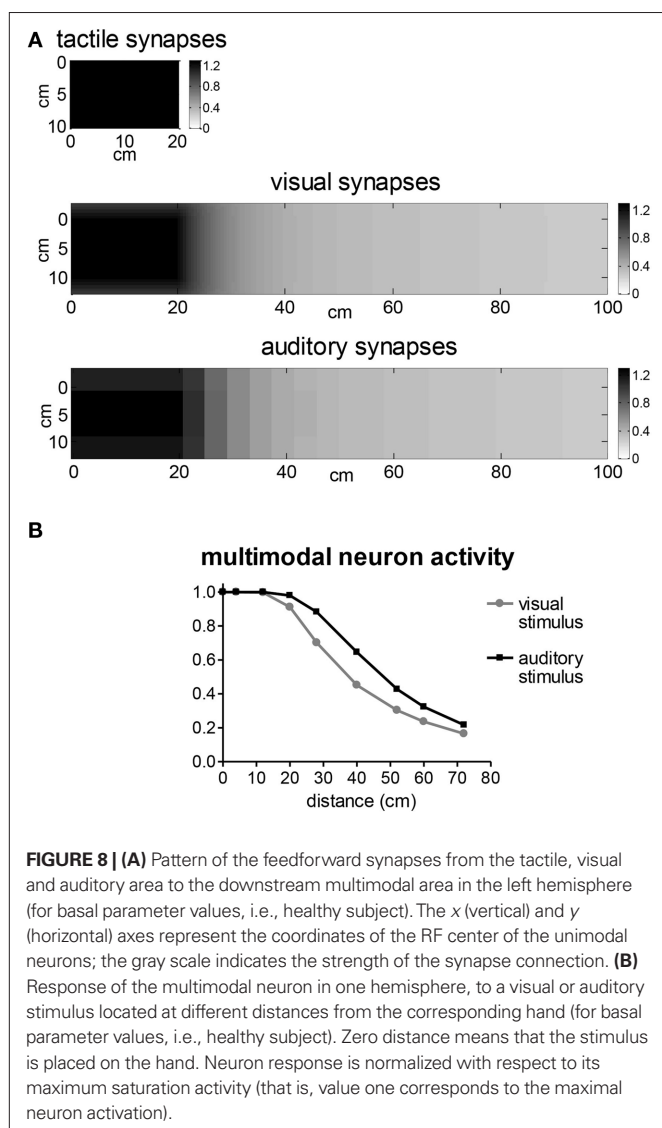
### Parameters of the neural network (healthy subject and RBD patient)

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 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).
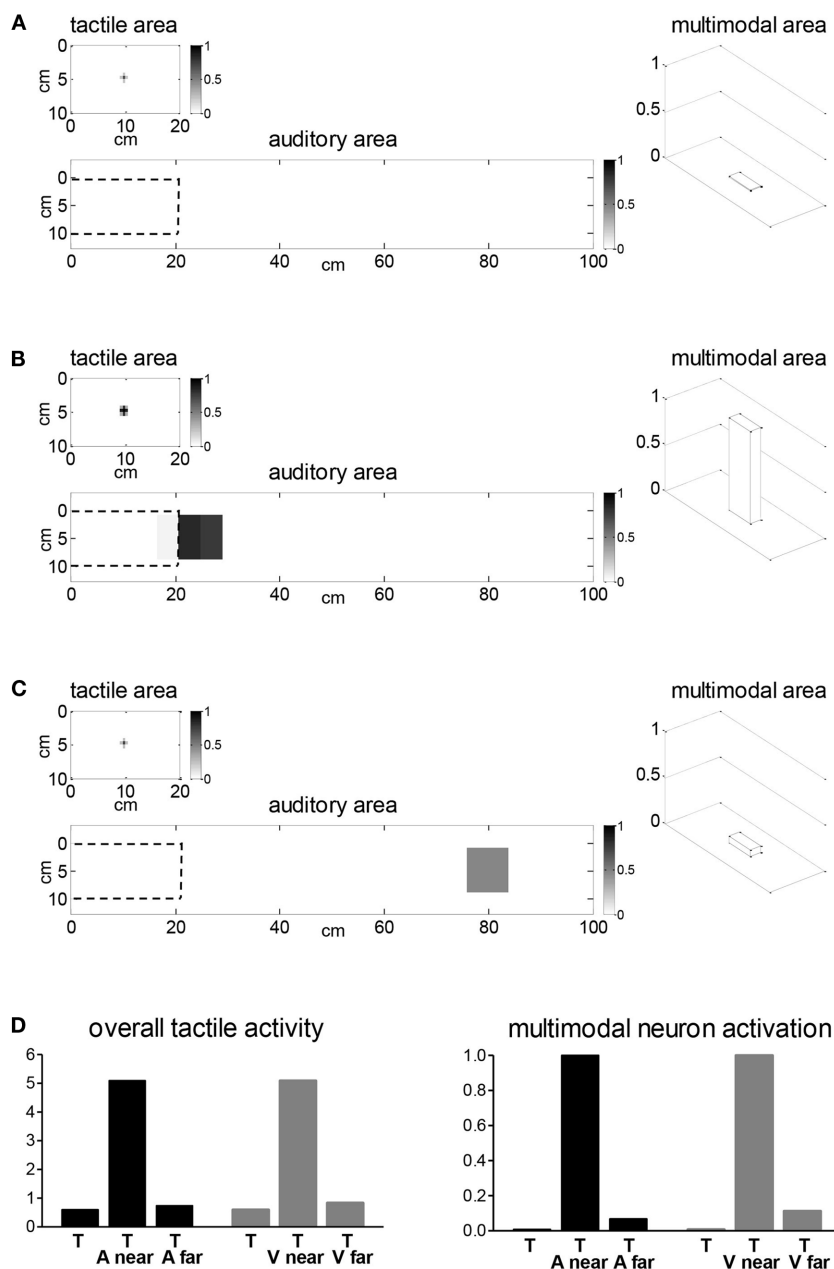
**FIGURE 9 | (A)** Network response to a weak tactile stimulus on the right hand. Plots show activity in tactile and in the auditory area (represented as gray plot) and in the multimodal area (represented via a 3D bar) of the left hemisphere. The dashed border within the auditory area delimits the auditory 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 mul-
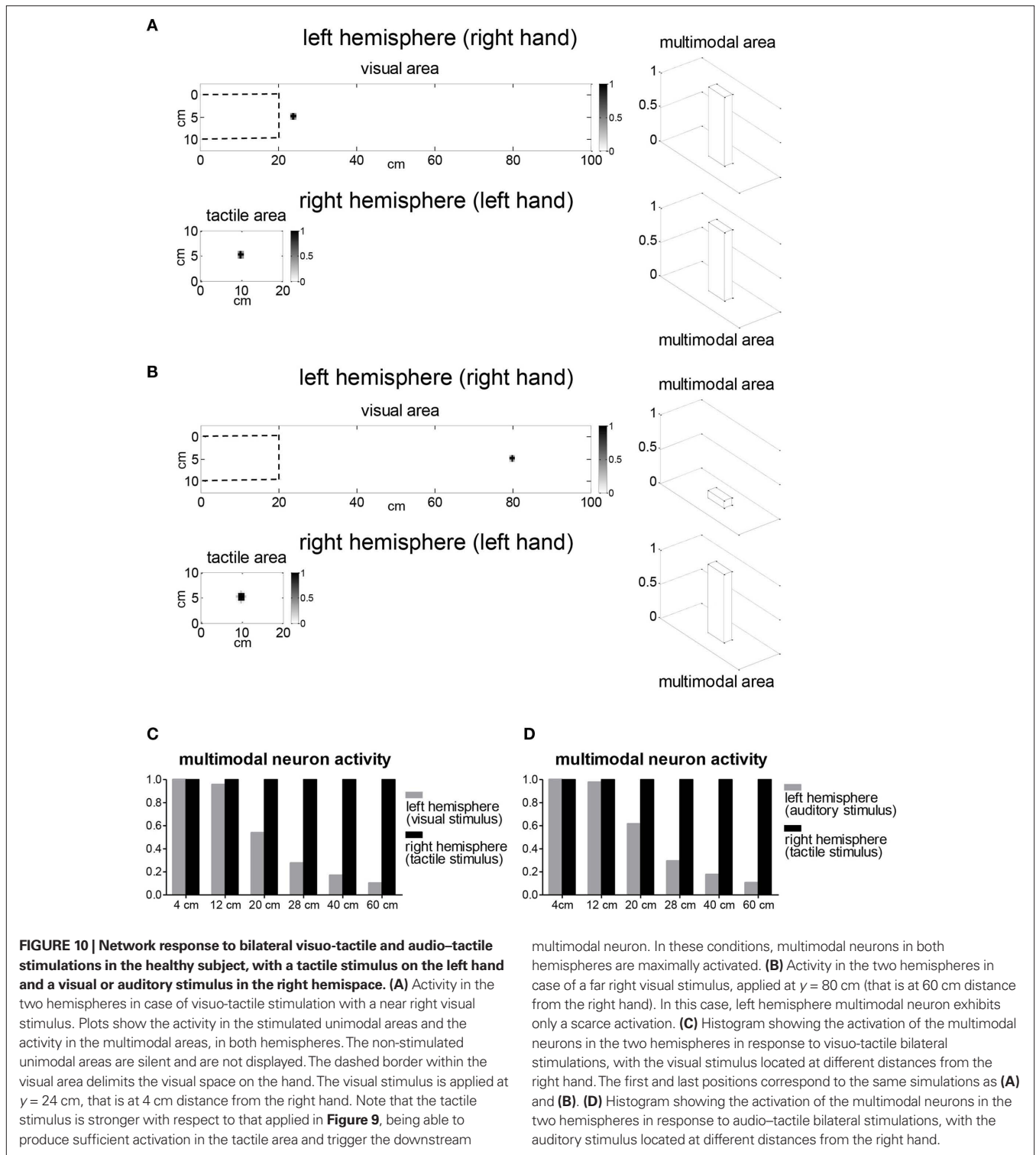
timodal neuron at its maximum level. The model predicts similar results in case of bilateral audio–tactile stimulations, with the auditory input at different distances from the right hand (**Figure 10D**).

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; Farné 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 dashed border within the visual area delimits the visual space on the hand. The dashed border within the visual area delimits the visual space on the hand. Note that the tactile 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
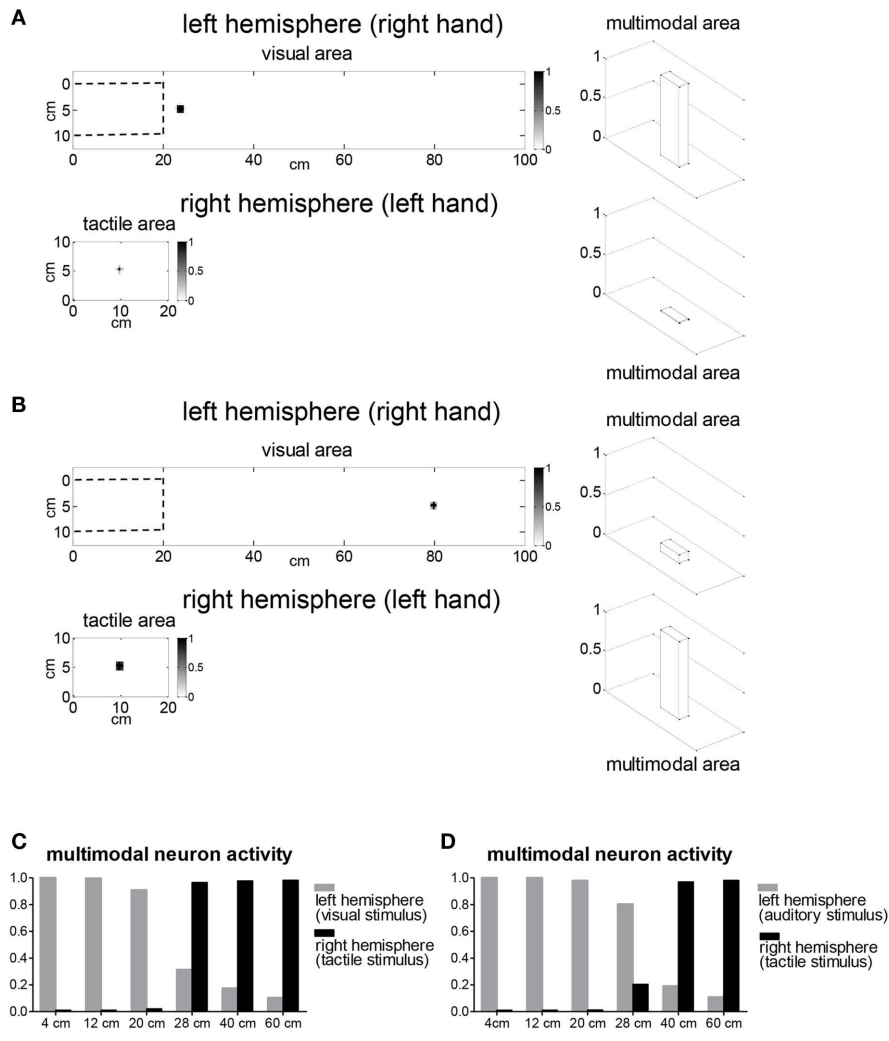
**FIGURE 11 | Network response to bilateral visuo-tactile and audio–tactile stimulations in the RBD patient, with a tactile stimulus on the left hand and a visual or auditory stimulus in the right hemispace.** Stimuli intensity is the same as in **Figure 10**. **(A)** Activity in the two hemispheres 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.
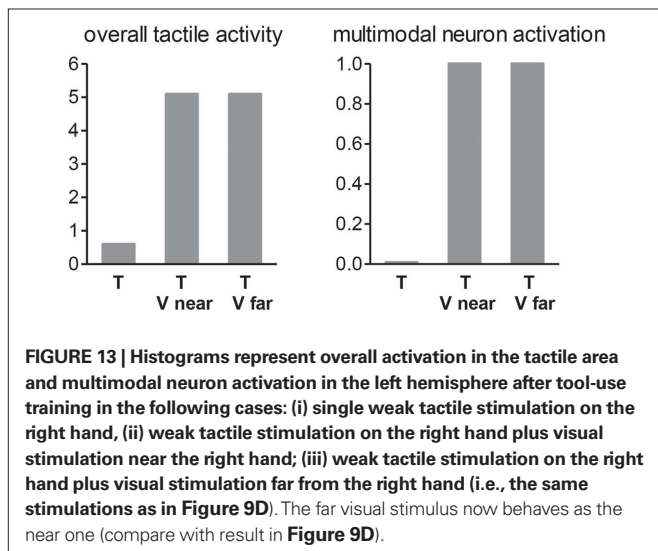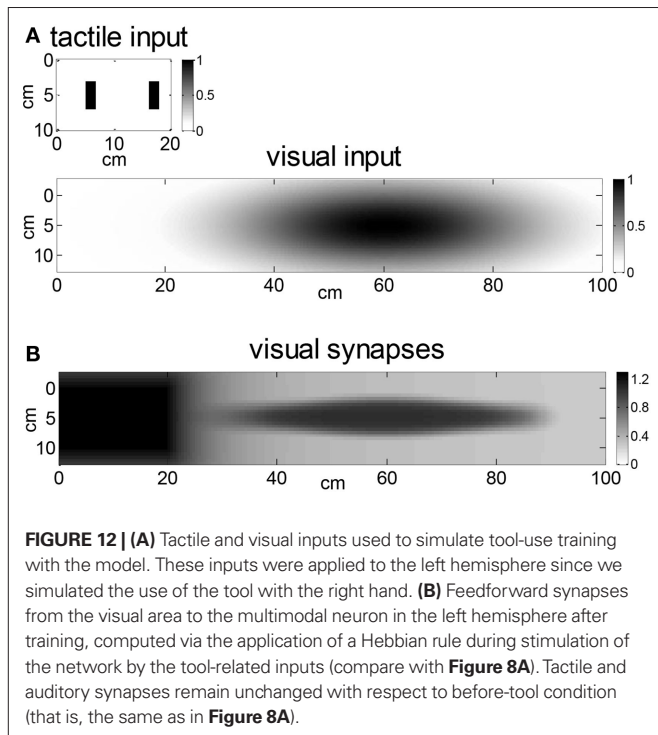
All equations and parameters concerning model training and plasticity can be found in our previous paper (Magosso et al., 2010a).

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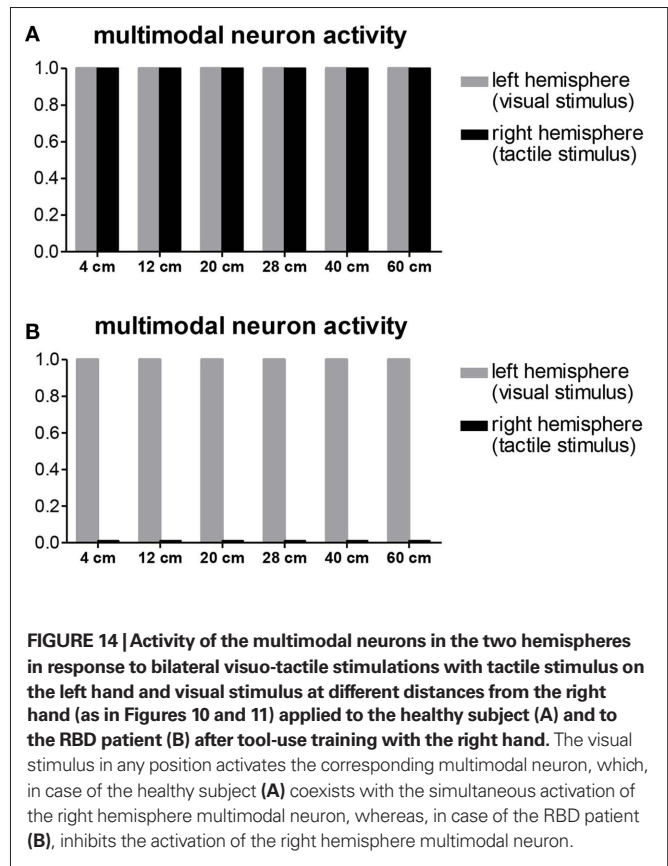
**Figure 14** reports the results of bilateral visuo-tactile stimulations, with the visual stimulus located at different distances from the right hand, in the healthy subject (**Figure 14A**) and in RBD patient (**Figure 14B**), after tool-use training. In the trained condition, the right visual stimulus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

Farné, 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). Contraleteral and ipsileteral 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.