

# The contribution of response conflict, multisensory integration, and body-mediated attention to the crossmodal congruency effect

Francesco Marini<sup>1,3</sup>  · Daniele Romano<sup>1,2</sup> · Angelo Maravita<sup>1,2</sup>

Received: 25 May 2016 / Accepted: 27 November 2016 / Published online: 2 December 2016  
© Springer-Verlag Berlin Heidelberg 2016

**Abstract** The crossmodal congruency task is a consolidated paradigm for investigating interactions between vision and touch. In this task, participants judge the elevation of a tactile target stimulus while ignoring a visual distracter stimulus that may occur at a congruent or incongruent elevation, thus engendering a measure of visuo-tactile interference (crossmodal congruency effect, CCE). The CCE reflects perceptual, attentional, and response-related factors, but their respective roles and interactions have not been set out yet. In two experiments, we used the original version of the crossmodal congruency task as well as ad hoc manipulations of the experimental setting and of the participants' posture for characterizing the contributions of multisensory integration, body-mediated attention, and response conflict to the CCE. Results of the two experiments consistently showed that the largest amount of variance in the CCE is explained by the reciprocal elevation of visual and tactile stimuli. This finding is compatible with a major role of response conflict for the CCE. Weaker yet distinguishable contributions come from multisensory integration, observed in the absence of response conflict, and from hand-mediated attentional binding, observed with the

modified posture and in the presence of response conflict. Overall, this study informs the long-standing debate about mechanisms underlying the CCE by revealing that the visuo-tactile interference in this task is primarily due to the competition between opposite response tendencies, with an additional contribution of multisensory integration and hand-mediated attentional binding.

**Keywords** Crossmodal congruency effect · Bodily attention · Hand-mediated binding · Multisensory integration · Visuo-tactile · Response conflict

## Introduction

The tactile perception of a sensory stimulus very often co-occurs with the visual perception of the same stimulus. For example, when we receive a high five from a friend or when we notice that a ladybug is about to land on our hand, we generally direct our eyes toward the hand. In these situations, stimuli that determine a tactile sensation on our body are also perceived visually. The existence of close links between vision and touch is supported by a body of behavioral (Botvinick and Cohen 1998; Ernst and Banks 2002; Pavani et al. 2000; Spence et al. 2004a, b), neurophysiological (Graziano and Gross 1993; Wallace and Stein 1997), neuropsychological (Spence et al. 2001), and neuroimaging studies (Macaluso et al. 2000a, b, 2002a, b; Zimmer and Macaluso 2007) in both non-human primates and humans (see Spence et al. 2007 for an overview). The prominent connection between vision and touch contributes to create a robust perception of tactile and visual events that occur on the body surface and in its close spatial proximity.

The crossmodal congruency task (CCT), first introduced by Driver and Spence (1998a, b), has been extensively used

---

✉ Francesco Marini  
francesco.pd@gmail.com

Daniele Romano  
daniele.romano@unimib.it

Angelo Maravita  
angelo.maravita@unimib.it

<sup>1</sup> Department of Psychology, University of Milano-Bicocca, Milan, Italy

<sup>2</sup> Milan Center for Neuroscience, Milan, Italy

<sup>3</sup> Department of Psychology, University of Nevada Reno, 1664 N Virginia St, Reno, NV 89557, USA

to investigate body-related interactions between the visual and the tactile modalities. In the typical version of this task, subjects are required to place their forearms on a table and hold a foam block with their index and thumb finger of each hand. Each block embeds a pair of visual and a pair of tactile stimulators in such a way that a visual and a tactile stimulator are placed in close proximity with one another on the top and on the bottom aspect of each block. On each trial, participants' task is to indicate the elevation (high or low) of the stimulus in one modality, typically touch, while ignoring a simultaneous distracter stimulus in the other modality, typically vision (but see Walton and Spence 2004 for the opposite visual-target tactile-distracter association). Importantly, distracter stimuli can occur either at the same (congruent) or at a different (incongruent) elevation relative to the tactile stimulus. The typical result consists in faster responses when both the tactile and the visual stimuli occur at congruent, rather than incongruent, elevations (Pavani et al. 2000; Spence et al. 2004a). The slowing down of responses on incongruent (relative to congruent) trials has been termed *crossmodal congruency effect* (CCE). Thus, the CCE represents a measure of crossmodal interference between visual and tactile stimuli delivered in the proximity of the body. The CCE has been used in a multiplicity of research realms, including crossmodal exogenous spatial attention (Driver and Spence 1998a, b), multisensory interactions in peripersonal space (Maravita et al. 2003; Spence et al. 2007; van Elk et al. 2013), rubber hand illusion (Pavani et al. 2000; Zopf et al. 2010, 2013), tool use (Maravita et al. 2002; Holmes et al. 2002, 2007; Sengül et al. 2013), temporal processing (Shore et al. 2006), distracter suppression (Marini et al. 2013), and even to study the embodiment of a robotic prosthesis (Marini et al. 2014).

A typical additional finding is that the CCE is larger when both the tactile target stimulus and the visual distracter stimulus are presented on the same side (i.e., both stimuli delivered either to the left or to the right side), relative to when visual and tactile stimuli are presented on opposite sides (i.e., the tactile stimulus to the left side and the visual stimulus to the right side, or vice versa). This finding has been reported in CCE studies by showing the existence of a significant interaction between relative elevation (congruent/incongruent) and relative side (same/opposite) of the tactile and visual stimuli (e.g., Spence et al. 2004a). However, an individual and systematic analysis of reaction times (RTs) to all possible combinations of target and distracter pairs has not been conducted so far. Then, the question of whether the reduced CCE for opposite-side (relative to same-side) stimuli arises from slower responses to congruent stimulus/target combinations or faster responses to incongruent pairings, or both, has not been addressed yet. This distinction could imply potentially different mechanisms underlying the CCE, which have not been completely set out so far.

At least three different mechanisms have been proposed to contribute to the CCE: multisensory integration, exogenous attention, and response conflict (Driver and Spence 1998a, b; Maravita et al. 2003; Spence et al. 2004a, b; Shore et al. 2006; Forster and Pavone 2008; Holmes 2012). Within this theoretical debate, the overarching aim of this work is to characterize the mechanisms underlying the CCE. Experiment 1 will investigate the role of multisensory integration and will directly compare individual CCE conditions against each other, thus providing with a systematic comparison of individual CCE conditions that has not been conducted so far. Experiment 2 will investigate the role of attention and response conflict in the CCE, with a particular focus on the attentional modulations related to the placement of body parts such as the participants' hands (hand-mediated attentional binding).

## Experiment 1

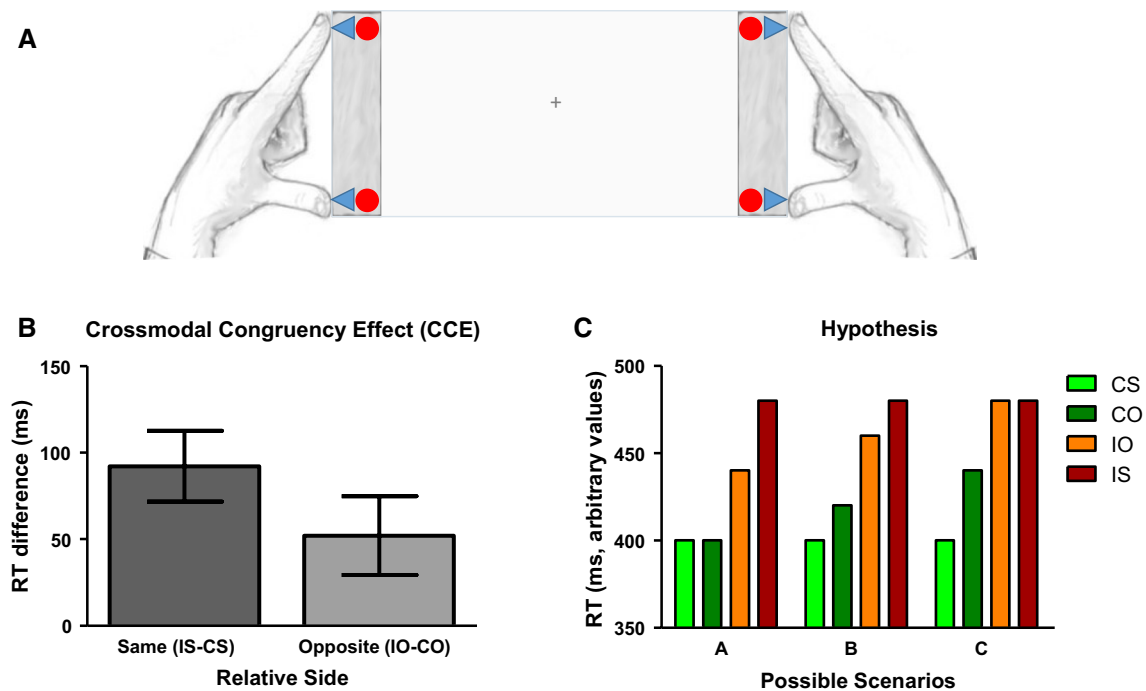
### Rationale and hypotheses

Multisensory integration refers to the combination of perceptual signals from different sensory modalities during their processing to form a unitary percept (Stein and Meredith 1993; Ernst and Bühlhoff 2004). In principle, multisensory integration could influence the CCE in two ways. First, because the integration is strongest when multisensory signals occur at the same location in space (Stein and Meredith 1993; Murray and Wallace 2011), multisensory integration could facilitate responses on congruent-same-side (vs. opposite-side) visuo-tactile pairs. Second, because localization errors are larger when two sensory modalities convey different spatial information (e.g., Alais and Burr 2004), more spatial ambiguity may occur when congruent visuo-tactile stimuli are presented on opposite (vs. same) sides. This could determine a slowing down of the responses on congruent-opposite-side versus same-side trials. Relatedly, qualitative reports of faster and more accurate responses on same- (vs. opposite-) side congruent trials are common in the CCE literature, but a systematical statistical comparison between the two conditions has not been reported so far. Experiment 1 was conducted in order to investigate the contribution of multisensory integration to the CCE while establishing the pattern of the CCE at the level of each individual condition.

### Methods

#### Participants

A previously unpublished dataset with a sample of 32 healthy volunteers (mean age  $\pm$  standard deviation: 24.9  $\pm$  5.2 years,



**Fig. 1** Experimental setup, rationale, and hypothesis. **a** The experimental setup of the crossmodal congruency task CCT), as used in previous experiments and in Experiment 1. Participants held one foam block with each hand and placed their index fingers on the upper side of each block and their thumbs on the lower side of each block, in correspondence with tactile stimulators (*blue triangles*) and visual stimulators (*red circles*). On each trial, one tactile stimulus and one visual stimulus were delivered at one of the four possible locations. Participants had to report the elevation of the tactile stimulus (high/low), regardless of its side, and ignore the visual stimulus. **b** The typical results of an experiment with the CCT. The crossmodal congruency effect (CCE) is calculated as the difference in reaction times (RT, *left axis*) between incongruent (visual and tactile stimuli presented at different elevations) and congruent (visual and tactile stim-

uli presented at the same elevation) trials. The CCE is larger when the visual and the tactile stimuli are presented on the same side relative to when they are presented on opposite sides. *CS* (congruent elevation, same side), *CO* (congruent elevation, opposite side), *IS* (incongruent elevation, same side), *IO* (incongruent elevation, opposite side) (data from Spence et al. 2004a, b). **c** Three different scenarios may account for the pattern of results of **b**. The same CCE would be observed in either scenario. However, in Scenario A congruent visuo-tactile stimulus pairs have the same RT regardless of their relative side; in Scenario B, there is a gradation in RTs when progressing through the four conditions; in Scenario C, incongruent visuo-tactile stimulus pairs have the same RT regardless of their relative side. The aim of Experiment 1 is to determine which scenario takes place in the CCE (color figure online)

10 males, 31 self-reportedly right-handed) was used in Experiment 1. All participants had normal or corrected-to-normal vision, were naïve as to the purpose of the experiment, and voluntarily agreed to take part in the research. This study was approved by the ethical committee of the University of Milano-Bicocca and was conducted in accordance with the Declaration of Helsinki (World Medical Organization 1996).

### Stimuli

The experimental apparatus for Experiment 1 consisted of a black vertical panel in which two foam blocks ( $8 \times 4 \times 3$  cm) were fixed to the left and the right side of a central fixation point, at a lateral distance of 25 cm. Two tactile stimulators (custom-made electromagnetic solenoids, Heijo Electronics, Beckenham, UK; [www.heijo.com](http://www.heijo.com)) were embedded in each block, at the top and the

bottom of the lateral side of each block. Two visual stimulators (red light-emitting diodes, LEDs) were embedded in each block and located in close proximity to the tactile stimulators (Fig. 1).

The tactile and the visual stimuli consisted of three 30-ms single pulses interleaved with two 30-ms off-phases, resulting in a total duration of 150 ms for each stimulus. Visual stimulation led tactile stimulation by 30 ms. This stimulus-onset asynchrony served to compensate for the different latencies of the visual and of the tactile sensory inputs and has been used in previous studies with the same paradigm (Spence et al. 2004a; Heed et al. 2010; Marini et al. 2013). The tactile stimulus was the stimulus to which participants had to respond to (target), while the visual stimulus was an irrelevant stimulus, which participants had to ignore (distracter). Presentation and timing of tactile and visual stimuli were under computer control through a

custom-made I/O stimulator box and the E-Studio software (Psychology Software Tools, Inc., Pittsburgh, PA, [www.psychtoolbox.org](http://www.psychtoolbox.org)).

### Task

Participants sat in a dimly illuminated room in front of a table, at a distance of 57 cm from the central fixation point, and performed a tactile elevation discrimination task. In Experiment 1, participants placed their forearms on the table and held the foam blocks (one in each hand) by keeping their index fingers on the upper tactile stimulators and their thumbs on the lower tactile stimulators. On each trial, a tactile stimulus (target) and a visual stimulus (distracter) were delivered at one out of the four possible locations (upper position on the right block corresponding to the right index; lower position on the right block corresponding to the right thumb; upper position on the left block corresponding to the left index; and lower position on the left block corresponding to the left thumb). Distracters were equally likely to occur at congruent or incongruent elevations and at the same or different side relative to targets. Thus, every possible spatial combination of targets and distracters was delivered with the same probability. This design included four experimental combinations as regards the respective locations of targets and distracters. The distracter might be located at the same elevation and on the same side, relative to the target (congruent-same trial, CS); at the same elevation, but on the opposite side (congruent-opposite trial, CO); at a different elevation, yet on the same side (incongruent-same trial, IS); at a different elevation and on the opposite side (incongruent-opposite trial, IO).

Participants responded to the elevation of tactile targets (high/low), regardless of the stimulation side (left/right) and while ignoring visual distracters. Responses were delivered using two foot pedals placed one underneath the participants' toes and one below their heel. Participants raised their toes to respond "high" (i.e., target on their index finger) or their heel to respond "low" (i.e., target on their thumb). This foot pedal method has been used to collect responses in many previous studies with the CCT (e.g., Spence et al. 2004a, b; Heed et al. 2010). Measures of reaction times (RT) and error rates were collected. The total duration of the task was about 30 min.

### Analysis

Statistical analyses were conducted on RTs and error rates, separately. Methodologically, we kept these two measures separated rather than combining them in the inverse efficiency score, as some previous CCE studies did (e.g., Shore et al. 2006), because the specific aim of Experiment 1 was to categorize RT and error rate patterns independently.

RTs were selected to eliminate outliers, excluding all trials below values of 200 ms (anticipatory responses) as well as all trials exceeding three standard deviations above the mean computed separately for each experimental condition (late responses) (Ratcliff 1993). The crossmodal congruency effect (CCE) was computed as the RT difference on incongruent minus congruent distracter conditions (e.g., Spence et al. 2004a). Statistical analyses of RT were conducted with repeated measures analysis of variance (ANOVA) as implemented in the software Statistica 6.0 (Statsoft Inc.). Significant ANOVA interactions were explored with paired *t* tests corrected for the family-wise error rate (FWER) with the Holm–Bonferroni method (Holm 1979).

Error rates were transformed in logit values, fitted to a binomial distribution, and analyzed with generalized linear mixed-effect models (Jaeger 2008) using the *lme4* package (version 1.1-12) in R (R Core Team 2016). Model selection was performed as follows. First, a model with random factors Subject (i.e., each participant) and Trial (i.e., each trial number across participants) was implemented, and then the most parsimonious random-effect structure was chosen by eliminating each factor that did not significantly improve the model's fit. Then, a mixed-effect model was generated by adding fixed-effect factors to the chosen random-effect structure. After the inclusion of each fixed-effect factor, the resulting model was tested against the random-effect model and only fixed-effect factors that contributed to improve the model's fit were included in the final structure of the mixed-effect model. All model comparisons used the Chi-square test ( $\alpha = .01$ ). Statistics for fixed-effect contrasts were estimated using the *lmerTest* package (version 2.0-32) and are reported with *z* values and the corresponding *p* values (as returned by the *lmerTest*). When appropriate, post hoc tests were conducted with the *phia* package (version 0.2-1), and the relative Chi-square statistics, degrees of freedom (*df*), and *p* values are reported (as returned by *phia*).

For significant effects on both RT and error rates, we report the mean value and the 95% confidence intervals (CI).

### Results

**Reaction Time** A  $2 \times 2$  ANOVA was conducted with factors Elevation (congruent/incongruent) and Side (same/opposite) of the target–distracter pairs. The main factor Elevation was significant [ $F(1, 31) = 145.5, p < 0.001$ ], with congruent trials eliciting faster RTs than incongruent trials (mean RTs  $\pm$  95% CI  $492 \pm 25$  and  $582 \pm 31$  ms, respectively). The main factor Side was not significant [ $F(1, 31) < 0.01, p > 0.99$ ]. The interaction between Elevation and Side was significant [ $F(1, 31) = 33.2, p < 0.001$ ].

The RT difference on incongruent minus congruent trials, namely the CCE, was larger when the target and the distracter were presented on the same side (mean CCE  $\pm$  95% CI  $122 \pm 22$  ms) relative to when they were presented on opposite sides (mean CCE  $\pm$  95% CI  $58 \pm 13$  ms) [ $t(31) = 5.76$ ,  $p < 0.001$ ] (Fig. 2a, left). In order to identify the scenario (among those presented in Fig. 1) corresponding to the observed CCE pattern, the interaction between Elevation and Side was explored by running the following comparisons: congruent-same versus congruent-opposite trials; congruent-opposite versus incongruent-opposite trials; and incongruent-opposite versus incongruent-same trials. All these three comparisons yielded significant results. Congruent-same trials (mean RT  $\pm$  95% CI  $477 \pm 24$  ms) were faster than congruent-opposite trials (mean RT  $\pm$  95% CI  $508 \pm 26$  ms) [ $t(31) = 6.35$ ,  $p < 0.001$ ]. Congruent-opposite trials were faster than incongruent-opposite trials (mean RT  $\pm$  95% CI  $567 \pm 30$  ms) [ $t(31) = 8.37$ ,  $p < 0.001$ ]. Incongruent-opposite trials were faster than incongruent-same trials (mean RT  $\pm$  95% CI  $598 \pm 35$  ms) [ $t(31) = 3.90$ ,  $p < 0.001$ ]. Therefore, these results suggest that the typical finding of larger CCE on same-side versus opposite-side trials arises from a finely graded pattern of RTs (Fig. 2b, left).

**Error Rate** The best-fit model included one random-effect factor (Subject), the two fixed-effect factors Elevation (congruent/incongruent) and Side (same/opposite), and the fixed-effect interaction Elevation\*Side. The analysis revealed higher error rates on incongruent relative to congruent trials (mean error rates  $\pm$  95% CI  $17.4 \pm 3.7$  and  $3.0 \pm 1.2$ , respectively) ( $Z = 14.76$ ,  $p < 0.001$ ). No significant differences were observed between same-side and opposite-side trials ( $Z = 0.52$ ,  $p = 0.60$ ). The interaction between Elevation and Side was significant ( $Z = 2.47$ ,  $p = 0.01$ ), indicating that on incongruent trials (but not on congruent trials), same-side stimuli yielded less accurate performance relative to opposite-side stimuli (mean error rates  $\pm$  90% CI  $20.1 \pm 4.6$  and  $14.7 \pm 3.2\%$ ) ( $X^2 = 19.28$ ,  $p < 0.001$ ) (Fig. 2c, left).

## Discussion

In principle, at least three different patterns may underlie the typical observation of a larger CCE on same-side versus opposite-side trials (Fig. 1c). The typically found same-side/opposite-side modulation of the CCE may arise from a speeding-up of responses on congruent-same (relative to congruent-opposite) trials, a slowing down of responses on incongruent-same (relative to incongruent-opposite) trials, or both. Experiment 1 helped distinguishing between these possibilities. Results clearly point to a graded pattern of RT responses (compare Fig. 2b, left panel, with Scenario B in Fig. 1c). However, error rates did not show any statistically

significant same-side/opposite-side modulation on congruent trials, thus pointing to a different pattern (compare Fig. 2c, left panel, with Scenario A in Fig. 1c).

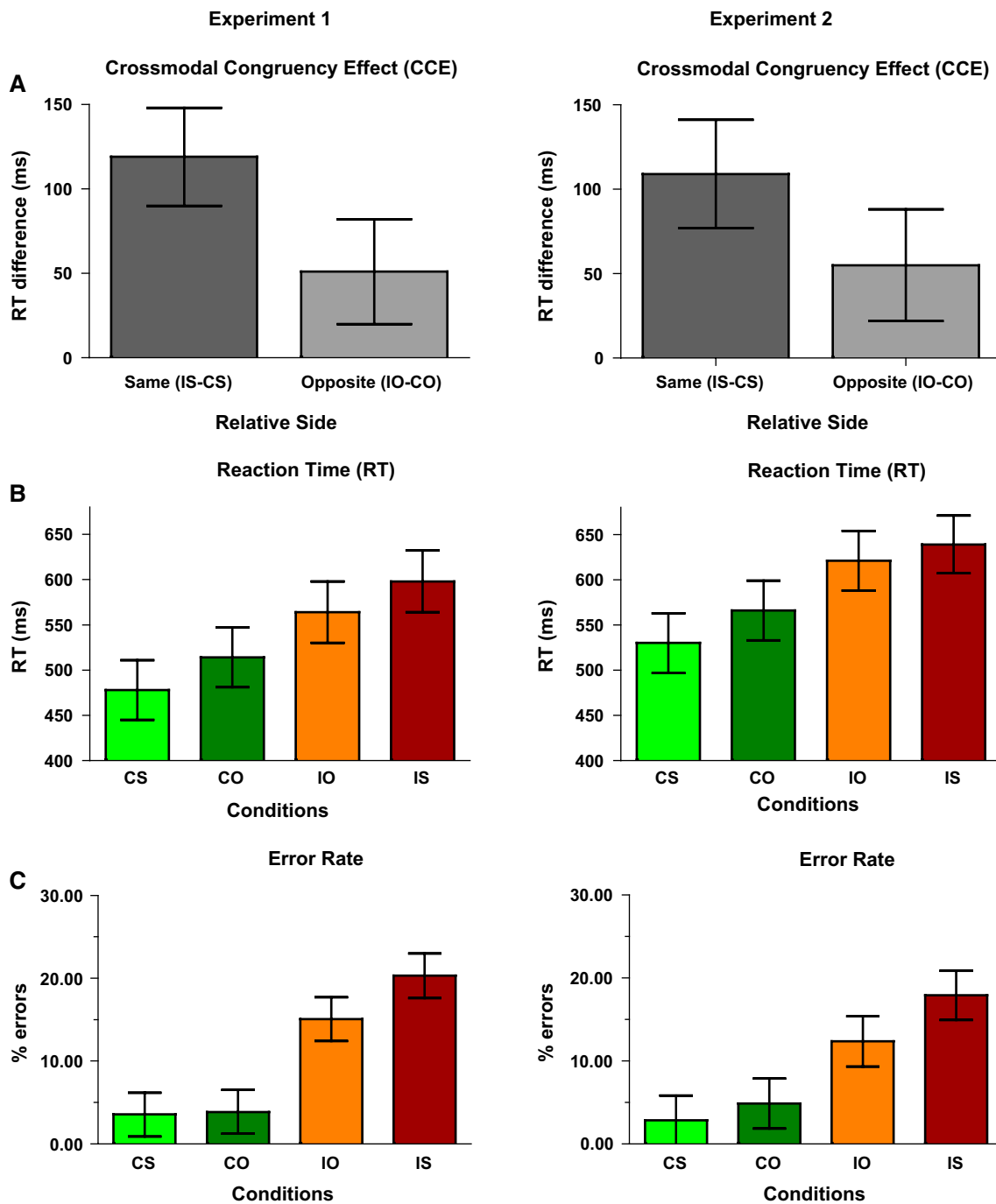
The direct comparison between congruent-same and congruent-opposite trials helped clarifying the role of multisensory integration in the CCE. For this, we focused on congruent trials because additional cognitive mechanisms—such as response conflict—do intervene on incongruent trials (Spence et al. 2004a; Forster and Pavone 2008). If we were to find a reliable statistical effect of faster and/or more accurate responses on congruent-same (vs. congruent-opposite) trials, this would indicate that multisensory integration plays a role in the crossmodal congruency task (at least on congruent trials). Such hypothesized difference was observed on RT data but not on error rates. The faster RTs observed on congruent-same (vs. congruent-opposite) trials might attest to visuo-tactile multisensory enhancement (Stein and Meredith 1993; Bolognini and Maravita 2007; Longo et al. 2012), although this seems rather limited because of the absence of significant differences in error rates. Furthermore, such a conclusion could be drawn more strongly if the speeding-up of responses to co-localized signals were observed relative to a “baseline” unimodal tactile condition, which was not included in the current study. Instead, a previous study that used a version of the CCE in which unimodal tactile trials were intermixed with crossmodal visuo-tactile trials found no difference in RTs between unimodal tactile trials and congruent visuo-tactile trials (Marini et al. 2013).

Overall, the contribution of multisensory integration to the CCE, although likely present to some extent, seems limited and other factors can possibly play a bigger role. Experiment 2 was conducted to distinguish the relative contributions of response conflict and body-mediated attention.

## Experiment 2

### Rationale and hypotheses

Visual events can capture covert spatial attention even when they should be ignored, causing slower RTs to a target stimulus when a salient yet task-irrelevant visual distracter is simultaneously presented (see Egeth and Yantis 1997 for review). In the CCT, salient visual distracters presented away from the target divert attention from the target location and may slow down responses to the target stimulus. However, whether or not attentional modulations of the CCT are influenced by the placement of body parts (such as the hands of the participants) has not been set out yet. When a body part is placed to connect two spatial locations, a special attentional binding may arise between the two locations, akin to the attentional facilitations observed



when two distinct visual stimuli are grouped with a line (Baylis and Driver 1992). We hypothesized that a form of attentional binding with object-based characteristics (e.g., Vecera and Farah 1994; Egly et al. 1994) may be involved in the CCE: the binding of spatial locations created by placing one's own hand across two spatial locations (*hand-mediated binding*). Experiment 2 assessed the impact of hand-mediated binding in the CCE with ad hoc manipulations of participants' posture.

Response conflict has been proposed as the major determinant of the CCE (Spence et al. 2004a; Forster and Pavone 2008). Response conflict refers to the involuntary activation of an inappropriate response representation in the stimulus–response mapping that may occur when a provoking yet task-irrelevant stimulus (or stimulus attribute) is presented simultaneously with the task-relevant stimulus, such as in the Flanker (Eriksen and Eriksen 1974), in the Stroop (1935), and in the Simon (1969) tasks. Active

**Fig. 2** Results of Experiment 1 and of Experiment 2. **a** *left and right* crossmodal congruency effect (CCE, in ms) measured as the average difference in reaction times (RT, in ms) between incongruent (visual and tactile stimuli presented at different elevations) and congruent (visual and tactile stimuli presented at the same elevation) trials for Experiment 1 (*left*) and for Experiment 2 (*right*). *Error bars* represent the standard error of the mean. The CCE is shown separately for same-side and for opposite-side visuo-tactile pairs. Results of Experiment 1 fully replicate previous findings (compare with Fig. 1, panel b). Results of Experiment 2 are substantially identical to those of Experiment 1, indicating that controlling both the distance between lateral stimulators and the “objecthood” of stimulator holders did not modify the CCE. **b** *left and right* average reaction times (RT, in ms) for the four experimental conditions, for Experiment 1 (*left*) and for Experiment 2 (*right*). *Error bars* represent the standard error of the mean. On congruent-same (CS) trials, visual and tactile stimuli were presented at the same elevation and on the same side; on congruent-opposite (CO) trials, visual and tactile stimuli were presented at the same elevation and on opposite sides; on incongruent-opposite (IO) trials, visual and tactile stimuli were presented at different elevations and on opposite sides; on incongruent-Same trials, visual and tactile stimuli were presented at different elevations and on the same side. This pattern of results supports Scenario B (Fig. 1, panel c). Results do not differ between Experiment 1 and Experiment 2. **c** *left and right* average error rates (percentage of errors) for the four experimental conditions (see above for the experimental conditions and relative acronyms) for Experiment 1 (*left*) and for Experiment 2 (*right*). *Error bars* represent the standard error of the mean. The pattern of results of error rates replicates the pattern of RT results in Experiment 2, while in Experiment 1 error rates did not differ between CS and CO trials

inhibition processes may successfully overcome this involuntary response tendency (Wijnen and Ridderinkhof 2007), although at a cost on reaction times, or may fail to do so, thus encompassing a cost in the overall performance accuracy (Logan and Cowan 1984; Mordkoff and Egeth 1993). In the CCT, the conflict between opposite response tendencies—the correct response primed by the tactile target and the incorrect response primed by the visual distracter—contributes the RT cost reflected by the CCE itself. However, since same-side and opposite-side incongruent distracters are equally conflicting with the target in terms of the required response, if the CCE were due uniquely to response conflict its magnitude should be similar for same-side and opposite-side pairs, which is not the case (e.g., Spence et al. 2004a). Therefore, response conflict does not suffice to explain the same-side versus different-side modulation of the CCE (see Holmes 2012 for a meta-analysis of CCE studies with tool use that focused on this aspect) and the postural manipulation of Experiment 2 helped to clarify why this is the case. We noted that in the classical CCT setup, same-side incongruent stimuli are always bound by the same hand while opposite-side stimuli are not. Therefore, we propose that the same-side versus opposite-side modulation of the CCE may be explained by hand-mediated attentional binding rather than by response conflict.

## Methods

### Participants

Eighteen new healthy volunteers (mean age  $\pm$  standard deviation:  $23.0 \pm 1.2$  years, 4 males, 17 self-reportedly right-handed) participated in Experiment 2. All participants had normal or corrected-to-normal vision, were naïve as to the purpose of the experiments, and voluntarily agreed to take part in the research. This study was approved by the ethical committee of the University of Milano-Bicocca and was conducted in accordance with the Declaration of Helsinki (World Medical Organization 1996).

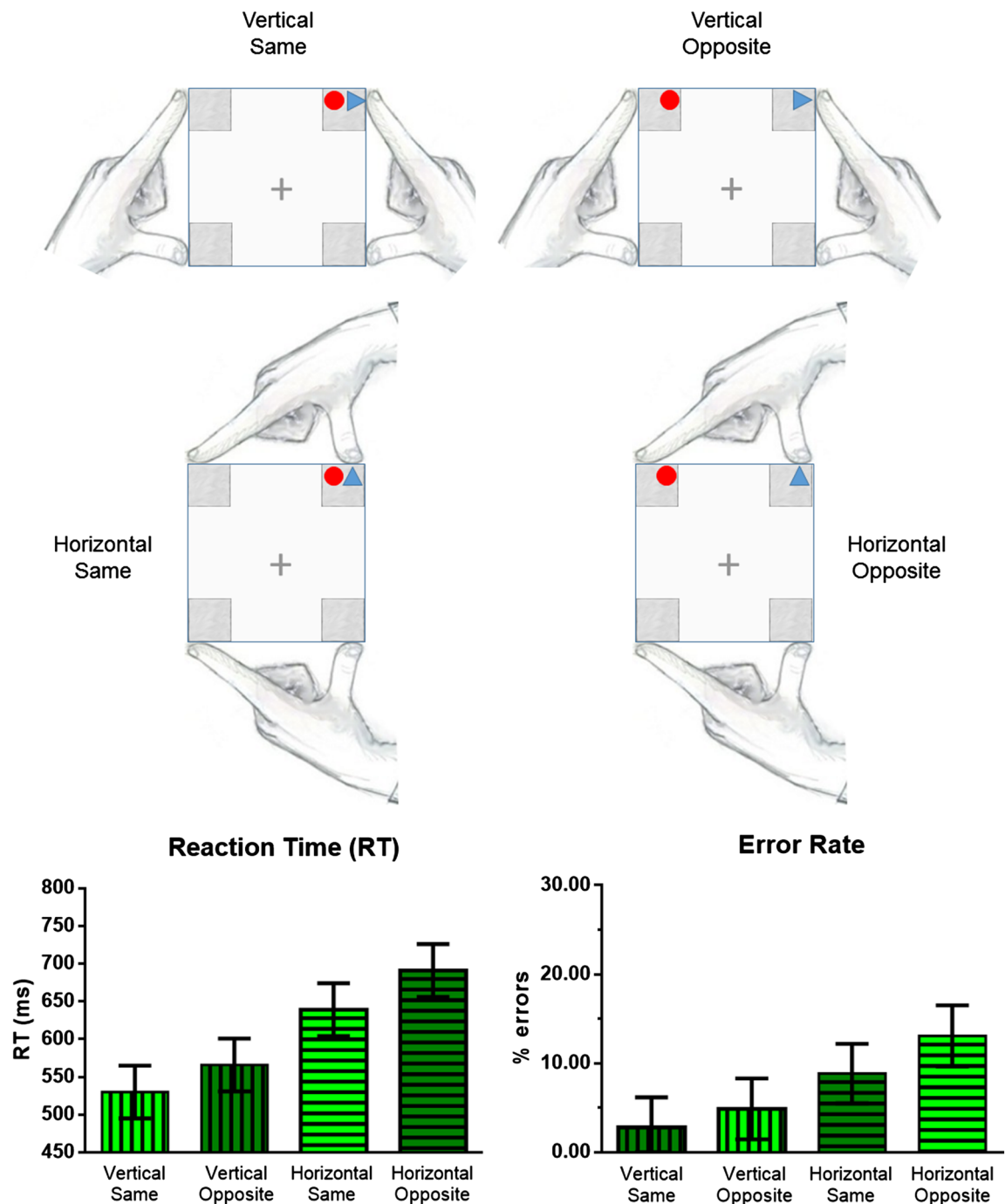
### Stimuli

The experimental apparatus for Experiment 2 consisted of a black vertical panel in which four small cubic foam blocks ( $3 \times 3 \times 3$  cm) were fixed, each one in correspondence with the vertex of an imaginary square centered in the fixation point. The resulting center-to-center distance between each block was 10 cm. One tactile stimulator (custom-made electromagnetic solenoids, Heijo Electronics, Beckenham, UK; [www.heijo.com](http://www.heijo.com)) and one visual stimulator (red light-emitting diodes, LED) were embedded in each block (see Figs. 3, 4). The tactile and the visual stimuli were identical to those used in Experiment 1, were delivered with the same stimulus-onset asynchrony (i.e., visual stimulus leading tactile stimulus by 30 ms), and were controlled with the same hardware and software used in Experiment 1.

### Task

Participants sat in front of a table at a distance of 57 cm from the central fixation point. Experiment 2 included two separate sessions that differed in the hands' posture (order counterbalanced across participants). In one session, similarly to Experiment 1, the index fingers were kept on the tactile stimulator of the upper blocks and the thumbs on the tactile stimulators of the lower blocks (“vertical” posture). In the other session, participants held the upper blocks with the index and thumb fingers of their left (or right) hand, and the lower blocks with the same fingers of the opposite hand (“horizontal” posture) (upper panels in Figs. 3, 4). Therefore, this postural manipulation created hand-mediated bindings that differed in the presence (vs. absence) of response conflict between the hand-bound locations (e.g., in the vertical posture, the binding was across blocks associated with conflicting responses, while in the horizontal posture the binding was across blocks associated with non-conflicting responses). Akin to Experiment 1, on each trial one tactile stimulus (target) and one visual stimulus

## Hand-mediated binding on Congruent Trials

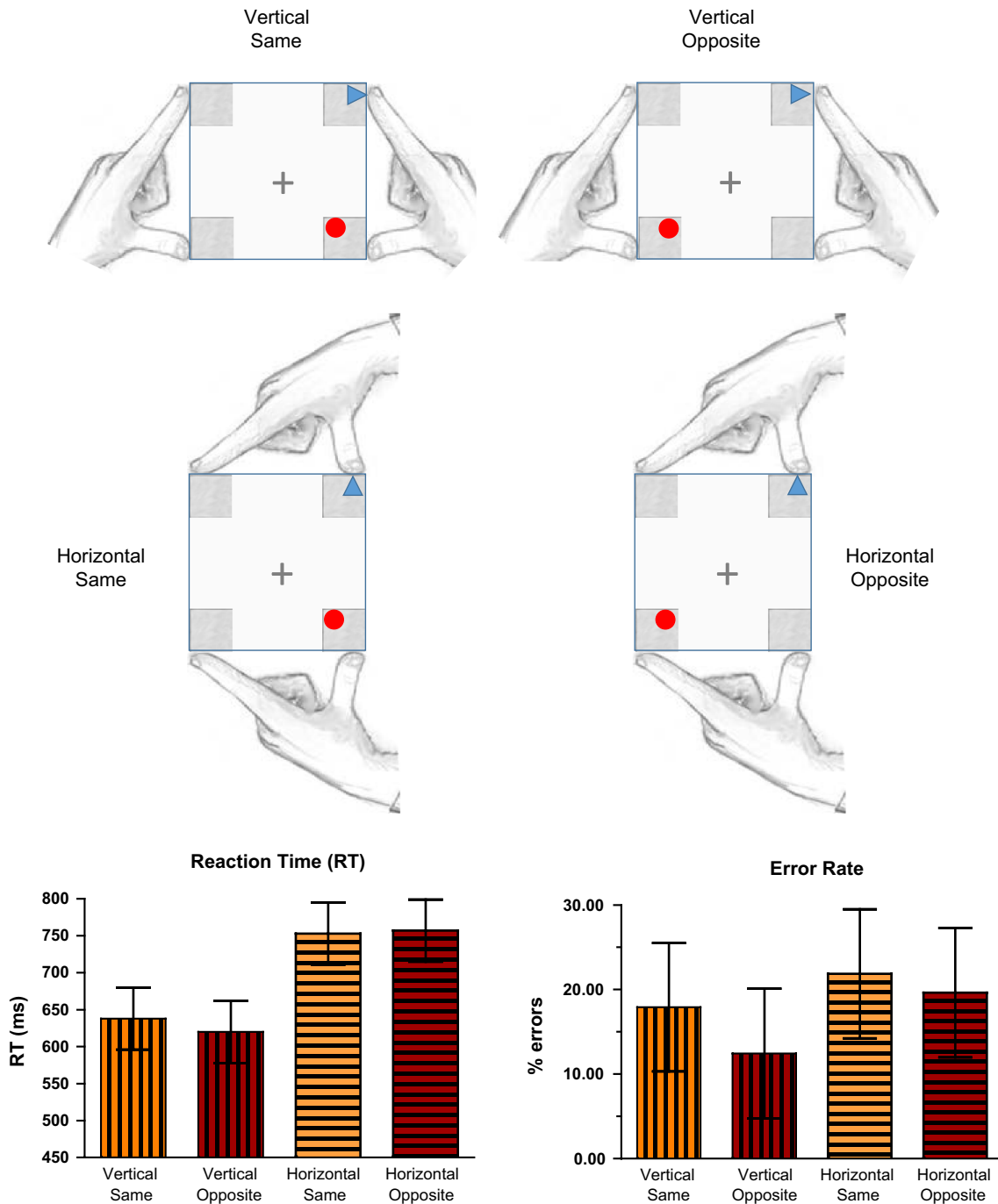


**Fig. 3** Results of the analysis of hand-mediated binding on congruent trials in Experiment 2 (corresponding to the conditions with no response conflict: RC−). *Upper panels* schematic representation of the experimental setup for congruent conditions of Experiment 2. *Blue triangles* represent tactile stimulators, and *red circles* represent visual stimulators. At each trial, one tactile stimulus and one visual stimulus are delivered at one of the four possible locations. Participants have to report the elevation of the tactile stimulus (high/low), regardless of its side (*left/right*) and hand position

(*vertical/horizontal*). For clarity, only the variant with the tactile stimulus in the *upper right* position is shown here. *Lower panels* average reaction times (in ms, *left graph*) and error rates (percentage of errors, *right graph*) for the congruent conditions of Experiment 2. *Error bars* represent the standard error of the mean. The vertical posture is associated with faster RTs and lower error rates relative to the horizontal posture, and same-side stimuli pairs are associated with faster RTs and lower error rates relative to opposite-side stimuli pairs (color figure online)



Hand-mediated binding on Incongruent Trials



**Fig. 4** Results of the analysis of hand-mediated binding on incongruent trials in Experiment 2 (corresponding to the conditions with response conflict: RC+). *Upper panels* schematic representation of the experimental setup for incongruent conditions of Experiment 2. *Blue triangles* represent tactile stimulators, and *red circles* represent visual stimulators. At each trial, one tactile stimulus and one visual stimulus are delivered at one of the four possible locations. Participants have to report the elevation of the tactile stimulus (high/low), regardless of its side (*left/right*) and hand position (*vertical/horizontal*). For clarity, only the variant with the tactile stimulus in the *upper right position* is shown here. *Lower panels* average

reaction times (in ms, *left graph*) and error rates (percentage of errors, *right graph*) for the incongruent conditions of Experiment 2. *Error bars* represent the standard error of the mean. The vertical posture is associated with faster RTs and lower error rates relative to the horizontal posture. An interaction is observed in RTs measures between posture and relative side of the visual stimulus: Same-side visual distracter stimuli are more interfering with the discrimination of the elevation of the tactile target stimulus, relative to opposite-side visual distracters, only when both stimuli are delivered to the same hand (i.e., in the vertical, relative to the horizontal, posture) (color figure online)

(distracter) were delivered to participants. The task was to indicate the elevation of the target (high/low in the absolute, space-based reference frame), regardless of its side, while ignoring the distracter. Response collection used the same foot pedal method described in Experiment 1 (paragraph 2.3). The total duration of the task was about 40 min.

### Analysis

Data analysis was conducted with the same methods described for Experiment 1 (paragraph 2.4). Briefly, a classical CCE analysis was conducted on trials with participants in the traditional “vertical” posture (note that, for comparison with Experiment 1 and with the existing CCE literature, data from the novel “horizontal” posture were not included in this analysis). Moreover, a global analysis including all trials was performed in order to investigate the contribution of hand-mediated binding to the same-side/opposite-side modulation of the CCE. However, because the same-side/opposite-side modulation has opposite signs on congruent and incongruent trials (i.e., RT differences on same minus opposite trials are negative- and positive-signed on congruent and incongruent trials, respectively), hand-mediated binding might be conditional on the presence of response conflict and therefore observable under incongruent conditions only (Holmes 2012). Therefore, in addition to the global analysis, exploratory analyses of hand-mediated binding were performed separately for conditions with and without response conflict (corresponding to incongruent and congruent trials, respectively).

## Results

### Analysis of the crossmodal congruency effect

**Reaction Time** One  $2 \times 2$  ANOVA was conducted on trials with the vertical posture with factors Elevation (congruent/incongruent) and Side (same/opposite) of the target–distracter pairs. A significant main effect of Elevation emerged, with faster responses on congruent versus incongruent trials (mean RTs  $\pm$  95% CI 551  $\pm$  25 and 633  $\pm$  32 ms, respectively) [ $F(1, 17) = 137.98, p < 0.001$ ] (Fig. 2a, right). Additionally, a significant main effect of Side was observed, with faster responses overall on same-side versus opposite-side trials (mean RTs  $\pm$  95% CI 588  $\pm$  29 and 596  $\pm$  27 ms, respectively) [ $F(1, 17) = 7.46, p = 0.01$ ]. More interestingly, a significant interaction between Elevation and Side was found [ $F(1, 17) = 24.58, p < 0.001$ ]. The RT difference between incongruent and congruent trials, namely the CCE, was larger when the target and the distracter were presented on the same side (mean CCE  $\pm$  95% CI 53  $\pm$  10 ms) relative to when they were presented on opposite sides (mean CCE  $\pm$  95% CI

11  $\pm$  24 ms) [ $t(17) = 4.96, p < 0.001$ ] (Fig. 2a, right). Further exploration of this interaction revealed that participants responded faster on same-side congruent trials relative to opposite-side congruent trials (mean RTs  $\pm$  95% CI 532  $\pm$  24 and 570  $\pm$  28 ms, respectively) [ $t(17) = 6.66, p < 0.001$ ]. On incongruent trials, however, participants responded faster on opposite-side (relative to same-side) trials (mean RTs  $\pm$  95% CI 623  $\pm$  27 and 643  $\pm$  38 ms, respectively) [ $t(17) = 2.69, p = 0.02$ ] (Fig. 2b, right). This pattern replicates the results of Experiment 1 and thus further supports the scenario (among those presented in Fig. 1c) of a finely graded pattern of RTs in the CCE.

**Error Rate** The best-fit model included one random-effect factor (Subject), the two fixed-effect factors Elevation (congruent/incongruent) and Side (same/opposite), and the fixed-effect interaction Elevation\*Side. The analysis revealed higher error rates on incongruent trials relative to congruent trials (mean error rates  $\pm$  95% CI 14.7  $\pm$  3.8 and 3.5  $\pm$  1.5%, respectively) ( $Z = 10.25, p < 0.001$ ) and on same-side trials relative to opposite-side trials (mean error rates  $\pm$  95% CI 10  $\pm$  2.9 and 8.1  $\pm$  2.5%, respectively) ( $Z = 2.49, p = 0.01$ ). The interaction between Elevation and Side was significant ( $Z = 3.91, p < 0.001$ ). Post hoc tests revealed that on congruent trials performance was more accurate on same-side versus opposite-side trials (mean error rates  $\pm$  95% CI 2.6  $\pm$  1.5 and 4.4  $\pm$  1.8%, respectively) ( $X^2 = 6.20, p = 0.01$ ), while on incongruent trials performance was more accurate on opposite-side versus same-side trials (mean error rates  $\pm$  95% CI 11.9  $\pm$  3.6 and 17.5  $\pm$  4.5%, respectively) ( $X^2 = 13.08, p < 0.001$ ) (Fig. 2c, right). Unlike Experiment 1, a finely graded pattern of error rates was observed in Experiment 2, thus fully replicating RT results.

### Global analysis of hand-mediated binding

**Reaction Time** One  $2 \times 2 \times 2$  ANOVA was conducted with factors Posture (vertical/horizontal), Elevation (congruent/incongruent), and Side (same/opposite). A significant main effect was found for each factor: Posture [ $F(1, 17) = 60.85, p < .001$ ], indicating faster responses in the vertical versus horizontal posture (mean RTs  $\pm$  95% CI 592  $\pm$  36 and 711  $\pm$  55 ms, respectively); Elevation [ $F(1, 17) = 130.59, p < .001$ ], indicating faster responses on congruent versus incongruent trials (mean RTs  $\pm$  95% CI 608  $\pm$  39 and 694  $\pm$  46 ms, respectively); Side [ $F(1, 17) = 33.44, p < .001$ ], indicating faster responses on same-side versus opposite-side trials (mean RTs  $\pm$  95% CI 642  $\pm$  44 and 661  $\pm$  40 ms, respectively). Two significant interactions were found. The interaction between Elevation and Side [ $F(1, 17) = 24.34, p < .001$ ] indicated that across postures responses were faster on congruent-same versus congruent-opposite trials (mean RTs  $\pm$  95%

CI  $587 \pm 42$  and  $630 \pm 50$  ms, respectively) but not on incongruent-same versus incongruent-opposite trials (mean RTs  $\pm$  95% CI  $697 \pm 52$  and  $692 \pm 51$  ms, respectively) [ $t(17) = 7.21$ ,  $p < .001$ , and  $t(17) = .91$ ,  $p = .37$ , respectively]. The interaction between Posture and Side [ $F(1, 17) = 10.33$ ,  $p = .005$ ] indicated that across elevations responses were faster on same-side versus opposite-side trials in the horizontal posture (mean RTs  $\pm$  95% CI  $696 \pm 65$  and  $725 \pm 64$  ms, respectively) and less so—yet still significantly faster—in the vertical posture (mean RTs  $\pm$  95% CI  $588 \pm 44$  and  $596 \pm 40$  ms) [ $t(17) = 5.17$ ,  $p < .001$ , and  $t(17) = 2.73$ ,  $p = .01$ , respectively]. Finally, the three-way interaction between Posture, Elevation, and Side was not significant [ $F(1, 17) = 1.27$ ,  $p = .27$ ]. We acknowledge that the lack of a significant three-way interaction allows to perform subsequent analyses (such as the two separate ANOVAS on congruent and on incongruent trials, respectively) with exploratory purposes only. Therefore, we advise that the related results should be taken with caution. Nonetheless, since we were interested in exploring the effects of posture separately for conflict and non-conflict trials (“Analysis” section), we conducted subsequent exploratory analyses of hand-mediated binding within congruent and incongruent trials (“Exploratory analysis of hand-mediated binding on congruent trials,” “Exploratory analysis of hand-mediated binding on incongruent trials” sections, respectively).

**Error Rate** The best fitting model included the random-effect factors Subject and Trial, the fixed-effect factors Elevation (congruent/incongruent), Side (same/opposite), and Posture (horizontal/vertical), and the fixed-effect interactions Elevation\*Side and Elevation\*Posture. All the remaining interactions did not improve the model’s fit and therefore were not included in the model. Fewer errors were observed on congruent trials relative to incongruent trials (mean error rates  $\pm$  95% CI  $7 \pm 1.7$  and  $17.6 \pm 3.4\%$ , respectively) ( $Z = 12.7$ ,  $p < 0.001$ ), on opposite-side versus same-side trials (mean error rates  $\pm$  95% CI  $11.9 \pm 3.5$  and  $12.6 \pm 2.8\%$ , respectively) ( $Z = 4.04$ ,  $p < 0.001$ ), and in the vertical versus horizontal posture (mean error rates  $\pm$  95% CI  $9.1 \pm 3.1$  and  $11.6 \pm 3.4\%$ , respectively) ( $Z = 8.62$ ,  $p < 0.001$ ). The interaction between Elevation and Side was significant ( $Z = 5.23$ ,  $p < 0.001$ ), indicating that on congruent trials performance was more accurate with same-side versus opposite-side stimuli (mean error rates  $\pm$  95% CI  $5.6 \pm 1.6$  and  $8.4 \pm 2\%$ , respectively) ( $X^2 = 14.44$ ,  $p < 0.001$ ) while on incongruent trials performance was more accurate with opposite-side versus same-side stimuli (mean error rates  $\pm$  95% CI  $15.5 \pm 3.8$  and  $19.6 \pm 3.6\%$ , respectively) ( $X^2 = 12.703$ ,  $df = 1$ ,  $p < 0.001$ ) (Fig. 2c, right). The interaction between Elevation and Posture was significant ( $Z = 4.83$ ,  $p < 0.001$ ), indicating that in the vertical posture performance was more accurate on congruent

versus incongruent trials (mean error rates  $\pm$  95% CI  $3.5 \pm 0.9$  and  $14.7 \pm 2.5\%$ , respectively) ( $X^2 = 138.67$ ,  $df = 1$ ,  $p < 0.001$ ). This was also the case in the horizontal posture ( $X^2 = 79.91$ ,  $df = 1$ ,  $p < 0.001$ ), yet the difference in error rates between congruent and incongruent trials was smaller (mean error rates  $\pm$  95% CI  $10.5 \pm 2.4$  and  $20.5 \pm 3.8\%$ , respectively).

#### *Exploratory analysis of hand-mediated binding on congruent trials*

**Reaction Time** One  $2 \times 2$  ANOVA was conducted on congruent trials factoring Posture (vertical/horizontal) and Side (same/opposite) of the target-distracter pairs. Both main effects of Posture and Side were significant, while the interaction between Posture and Side was not significant [ $F(1, 17) = 1.49$ ,  $p = 0.24$ ] (lower left graph in Fig. 3). Responses were overall faster in the vertical (mean RT  $\pm$  95% CI  $551 \pm 25$  ms) relative to the horizontal posture (mean RT  $\pm$  95% CI  $666 \pm 41$  ms) [ $F(1, 17) = 63.78$ ,  $p < 0.001$ ]. Responses were also faster on same-side (mean RT  $\pm$  95% CI  $587 \pm 29$  ms) relative to opposite-side trials (mean RT  $\pm$  95% CI  $630 \pm 34$  ms) [ $F(1, 17) = 53.46$ ,  $p < 0.001$ ].

**Error Rate** The best fitting model included the random-effect factors Subject and Trials, both the fixed-effect factors Posture (vertical/horizontal) and Side (same/opposite), but not the fixed-effect interaction Posture\*Side. Error rates were smaller in the vertical posture relative to the horizontal posture (mean error rates  $\pm$  90% CI  $3.5 \pm 1.5$  and  $10.5 \pm 2.6\%$ , respectively) ( $Z = 8.58$ ,  $p < 0.001$ ). Moreover, error rates were smaller on same-side versus opposite-side trials (mean error rates  $\pm$  95% CI  $5.6 \pm 2$  and  $8.4 \pm 2.1\%$ , respectively) ( $Z = 4.08$ ,  $p < 0.001$ ) (lower right graph in Fig. 3).

#### *Exploratory analysis of hand-mediated binding on incongruent trials*

**Reaction Time** One  $2 \times 2$  ANOVA was conducted on incongruent trials factoring Posture (vertical/horizontal) and Side (same/opposite) of the target-distracter pairs. The main effect of Posture was significant, with the vertical posture eliciting faster responses (mean RT  $\pm$  95% CI  $633 \pm 32$  ms) relative to the horizontal posture (mean RT  $\pm$  95% CI  $755 \pm 47$  ms) [ $F(1, 17) = 48.68$ ,  $p < 0.001$ ]. The main effect of Side was not significant [ $F(1, 17) = 0.83$ ,  $p = 0.37$ ]. Interestingly, a significant interaction between Posture and Side was observed [ $F(1, 17) = 7.49$ ,  $p = 0.01$ ] (lower left graph in Fig. 4). Direct comparisons revealed that same-side trials elicited slower responses relative to opposite-side trials in the vertical posture (i.e., when same-side incongruent stimuli were

delivered to the same hand) [ $t(17) = 2.69, p = 0.02$ ], but no significant difference was observed between same-side and opposite-side trials in the horizontal posture (i.e., when same-side incongruent stimuli were delivered to different hands) [ $t(17) = 1.1, p = 0.29$ ]. Therefore, the same-side/opposite-side modulation of the CCE was statistically significant only when incongruent-same target–distracter pairs were subjected to hand-mediated binding.

**Error Rate** The best fitting model included two random-effect factors (Subject, Trial), the fixed-effect factors Posture (vertical/horizontal), Side (same/opposite), while the fixed-effect interaction Posture\*Side contributed only marginally ( $p = .07$ ) to the improvement of model fit and therefore was not included. Error rates were smaller in the vertical relative to the horizontal posture (mean error rates  $\pm$  95% CI  $14.7 \pm 3.8$  and  $20.5 \pm 4.6\%$ , respectively) ( $Z = 4.67, p < 0.001$ ). Moreover, error rates were higher on same-side trials relative to opposite-side trials (mean error rates  $\pm$  95% CI  $19.6 \pm 4.4$  and  $15.5 \pm 3.8\%$ , respectively) ( $Z = 3.36, p < 0.001$ ) (lower right graph in Fig. 4).

## Discussion

The standard analysis of the CCE in Experiment 2 replicated most results of Experiment 1. A graded pattern was observed in reaction time data (compare Fig. 2b, right panel, with Scenario B in Fig. 1c) as well as in error rates (compare Fig. 2c, right panel, with Scenario B in Fig. 1c). The latter finding represents a difference with respect to Experiment 1, in which no same-side/opposite-side modulation of error rates was observed on congruent trials. Interestingly, locations of opposite-side pairs of stimulators in Experiment 2 were closer in external space relative to Experiment 1 (see Methods). Although speculative, it is tempting to propose that the higher error rate difference observed on congruent-opposite (vs. congruent-same) trials of Experiment 2 might arise from a greater multisensory interference of the visual distracter, possibly due to its closer spatial proximity to the tactile target.

The experimental manipulation of participants' posture in Experiment 2 helped revealing the role of hand-mediated attentional binding in the CCE. Although the three-way interaction between Posture, Elevation, and Side was not significant, subsequent exploratory analyses indicated a significant difference between hand-bound and non-hand-bound conditions on incongruent trials, but not on congruent trials. This suggests that a one's body part, such as the hand, may be capable of creating a binding between separate locations in external space, possibly through exogenous mechanisms with object-based attention characteristics. However, results should be taken with caution, and further research is needed to investigate the contribution of hand-mediated binding to the increase of the crossmodal

interference between visual and tactile stimuli on incongruent trials.

## General discussion

This study investigated the contribution of response conflict, multisensory integration, and hand-mediated binding to the crossmodal congruency effect (CCE). The classical CCE analysis of Experiments 1 and 2 replicated the typically observed pattern of (1) a slowing down on incongruent (vs. congruent) trials; and (2) an interaction between the relative side and the relative elevation of tactile targets and visual distracters. Additional analyses contrasting individual experimental conditions, which had not been reported by previous studies, revealed a finely graded pattern of RT responses in both experiments, and a finely graded pattern of error rates in Experiment 2.

Our results help characterizing the role of multisensory integration in the CCE. In principle, the advantage on the speeded elevation discrimination of a tactile stimulus when it is co-localized with a visual distracter, relative to when the visual distracter is presented at the same elevation but farther apart, may indicate that the visual signal, despite being irrelevant, actually facilitates performance, possibly attesting to the multisensory integration of visual and tactile signals (Stein and Meredith 1993). However, the speeding-up of responses to co-localized signals observed here was relative to non-co-localized signals rather than to some "baseline" condition. A recent work showed that RTs to tactile stimuli on congruent-same trials were no different in the absence or in the presence of a visual distracter stimulus (Marini et al. 2013), thus apparently contradicting the occurrence of crossmodal enhancements (Bolognini and Maravita 2007; Longo et al. 2012) in the CCT. Alternatively, the RT difference on congruent-same versus congruent-opposite trials might be interpreted as a relative slowing down (relative to some unimodal baseline condition) when a visual distracter is presented far apart from the target. This would support a different role for multisensory integration in the CCE, not in terms of facilitation on congruent-same trials, but rather in terms of crossmodal interference leading to the slowing down on congruent-opposite (relative to congruent-same) trials. If this interpretation is correct, then the slowing down might be ascribed to greater spatial uncertainty about target location in the presence of opposite-side distracters. Greater spatial uncertainty about target location appears in line with the observation of larger error rates on congruent-opposite relative to congruent-same trials in Experiment 2. Moreover, target mislocalization in the CCE has been documented by previous work, albeit comprising a small sample size ( $n = 5$ ), with an unspeeded version of the task (see Appendix in Spence

et al. 2004a) in which the tactile target stimulus was ventriloquized by a visual distracter stimulus.

The decreased performance (slower RTs and higher error rates) on incongruent (vs. congruent) trials, which was consistently observed in both experiments, fully complies with the predominant role of response conflict in the CCE, an idea that is well supported by the existing literature (Spence et al. 2004a; Forster and Pavone 2008). Based on three behavioral variations of the CCT, Spence and colleagues observed that “perceptual mislocalization may account for only a small component of CCE” and therefore “the crossmodal congruency effect is likely to primarily reflect response competition” (Spence et al. 2004a). Moreover, psychophysiological modulations of conflict-related ERP components have been isolated in the CCT, with increased frontocentral N2 for incongruent (vs. congruent) trials and larger error-related negativity (ERN) subsequent to errors on congruent (vs. incongruent) trials (Forster and Pavone 2008). This pattern of ERP responses is consistent with the response conflict account of the CCE. Our data fully agree with the aforementioned studies on the major role of response conflict in the CCE because in both experiments, we observed a significant slowing down of performance on incongruent versus congruent trials, with both same-side and opposite-side visuo-tactile stimuli, both in the vertical and in the horizontal posture.

Experiment 2 brought some methodological advancements to the setup of the CCT. In most previous CCE studies (but see Pavani and Castiello 2004), the binding between tactile and visual stimuli on each side was not only created by the participants’ hand, since the two stimulators were embedded in the same physical object (i.e., a foam block). Therefore, the larger CCE on incongruent-same relative to incongruent-opposite trials might be partially related to the physical properties of the block. In Experiment 2, we controlled this specific aspect by placing each stimulator in a separate physical object (i.e., a smaller foam block) and results confirmed the same gradual pattern of effects already found in Experiment 1. This replication under more controlled conditions ensures that the pattern of CCE does not depend on (and is not modulated by) the objecthood of the foam blocks embedding the two pairs visuo-tactile stimulators on each side.

A previously uninvestigated issue that was considered in our study regards the role in the CCE of body-mediated mechanisms such as attentional binding induced by the hand posture, or hand-mediated binding. This was investigated with the posture manipulation of Experiment 2. First, a large yet unanticipated effect of posture on the CCE was observed on both RTs and error rates. Participants responded slower and made more errors when adopting the horizontal versus the vertical posture. Effects of posture on cognitive performance have been described in the literature

(see Vercruyssen and Simonton 1994 for review). In our study, participants anecdotally reported that the horizontal position was less “natural” and this may reflect the fact that a larger, sustained effort was required for the maintenance of the horizontal (relative to the vertical) posture. Interestingly, subjects performing cognitive tasks while maintaining body positions that are less stable and physically effortful to maintain manifest a decrease in performance, which has been related to the increased need for cognitive attentional control (Teasdale et al. 1993). Therefore, the RT slowing and the error rate increase in the horizontal (vs. vertical) posture may be related to the necessity of maintaining a posture that requires greater effort to control.

Exploratory comparisons on the effects of hand-mediated binding with and without response conflict allowed us to make some further considerations. Without response conflict, responses were faster when target and distracter stimuli occurred on the same (vs. opposite) side, regardless of posture. Therefore, without response conflict, the presence (or absence) of hand-mediated binding did not seem to modulate the same-side versus opposite-side difference in the CCE. With response conflict, however, hand-mediated binding modulated performance. In fact, on incongruent trials, visual distracters located on the same side of tactile targets yielded larger interference relative to opposite-side distracters but only when the participants’ hand was binding the locations of target and distracter stimuli (i.e., in the vertical but not in the horizontal posture). Therefore, hand-mediated binding might account for the typically found larger interference on incongruent-same relative to incongruent-opposite trials, an effect that cannot be explained by response conflict because it is present in both types of trial.

What mechanism(s) might be responsible for hand-mediated binding effects in the CCE? Possibly, hand-mediated binding may facilitate the formation of a crossmodal attentional object in the peripersonal space across which attention spreads out (Busse et al. 2005; Fiebelkorn et al. 2010; Talsma et al. 2010). Consequently, the selection of one between two competing response tendencies might be more difficult when such responses originate from within the same attentional object (Duncan 1984; Baylis and Driver 1992; Egly et al. 1994; Vecera and Farah 1994). Due to such inspiring previous evidence, it would be of interest to set up future studies to explore the role of hand-mediated binding in the CCE, given that the analyses provided in the present paper are only exploratory. Moreover, these effects were observed on incongruent trials only. This could set specific boundaries for the modulatory effects of any hand-mediated attentional binding, possibly indicating that hand-mediated binding does not *necessarily* create a unitary attentional object with competing attentional effects. Rather, in the CCE these competing attentional effects seem to emerge only when opposite response tendencies

also compete within the attentional object created by hand-mediated binding.

## Conclusion

This study provided a theoretical and empirical characterization of visuo-tactile interference effects typically observed in the crossmodal congruency task. We propose a multifactorial interpretation of the CCE in terms of three contributing mechanisms: multisensory integration, hand-mediated attentional binding, and response conflict. Multisensory integration seems involved in the modulation the same-side/opposite-side difference in the absence of conflict, while hand-mediated attentional binding contributes to the same-side/opposite-side modulation in the presence of conflict. Response conflict represents the major determinant of the CCE.

## Compliance with ethical standards

**Conflict of interest** The authors declare no competing financial interests.

## References

- Alais D, Burr D (2004) The ventriloquist effect results from near-optimal bimodal integration. *Curr Biol* 14(3):257–262
- Baylis GC, Driver J (1992) Visual parsing and response competition: the effect of grouping factors. *Percept Psychophys* 51(2):145–162
- Bolognini N, Maravita A (2007) Proprioceptive alignment of visual and somatosensory maps in the posterior parietal cortex. *Curr Biol* 17(21):1890–1895
- Botvinick M, Cohen J (1998) Rubber hands ‘feel’ touch that eyes see. *Nature* 391(6669):756
- Busse L, Roberts KC, Crist RE, Weissman DH, Woldorff MG (2005) The spread of attention across modalities and space in a multisensory object. *Proc Natl Acad Sci USA* 102(51):18751–18756
- Driver J, Spence C (1998a) Attention and the crossmodal construction of space. *Trends Cogn Sci* 2(7):254–262
- Driver J, Spence C (1998b) Crossmodal attention. *Curr Opin Neurobiol* 8:245–253
- Duncan J (1984) Selective attention and the organization of visual information. *J Exp Psychol Gen* 113(4):501
- Egeth HE, Yantis S (1997) Visual attention: control, representation, and time course. *Annu Rev Psychol* 48(1):269–297
- Egly R, Driver J, Rafal RD (1994) Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects. *J Exp Psychol Gen* 123(2):161
- Eriksen BA, Eriksen CW (1974) Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept Psychophys* 16(1):143–149
- Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415:429–433
- Ernst MO, Bühlhoff HH (2004) Merging the senses into a robust percept. *Trends Cogn Sci* 8(4):162–169
- Fiebelkorn IC, Foxe JJ, Molholm S (2010) Dual mechanisms for the cross-sensory spread of attention: how much do learned associations matter? *Cereb Cortex* 20(1):109–120
- Forster B, Pavone EF (2008) Electrophysiological correlates of crossmodal visual distractor congruency effects: evidence for response conflict. *Cogn Affect Behav Neurosci* 8:65–74
- Graziano MSA, Gross CG (1993) A bimodal map of space: somatosensory receptive fields in the macaque putamen with corresponding visual receptive fields. *Exp Brain Res* 97:96–109
- Heed T, Habets B, Sebanz N, Knoblich G (2010) Others’ actions reduce crossmodal integration in peripersonal space. *Curr Biol* 20(15):1345–1349
- Holm S (1979) A simple sequentially rejective multiple test procedure. *Scand J Stat* 6:65–70
- Holmes NP (2012) Does tool use extend peripersonal space? A review and re-analysis. *Exp Brain Res* 218(2):273–282
- Holmes NP, Calvert GA, Spence C (2002) Extending or projecting peripersonal space with tools? Multisensory interactions highlight only the distal and proximal ends of tools. *Neurosci Lett* 372:62–67
- Holmes NP, Calvert GA, Spence C (2007) Tool use changes multisensory interactions in seconds: evidence from the crossmodal congruency task. *Exp Brain Res* 183(4):465–476
- Jaeger TF (2008) Categorical data analysis: away from ANOVAs (transformation or not) and towards logit mixed models. *J Mem Lang* 59(4):434–446
- Logan GD, Cowan WB (1984) On the ability to inhibit thought and action: a theory of an act of control. *Psychol Rev* 91(3):295
- Longo MR, Musil JJ, Haggard P (2012) Visuo-tactile integration in personal space. *J Cogn Neurosci* 24(3):543–552
- Macaluso E, Frith CD, Driver J (2000a) Modulation of human visual cortex by crossmodal spatial attention. *Science* 289:1206–1208
- Macaluso E, Frith CD, Driver J (2000b) Selective spatial attention in vision and touch: unimodal and multimodal mechanisms revealed by PET. *J Neurophysiol* 83:3062–3075
- Macaluso E, Frith CD, Driver J (2002a) Crossmodal spatial influences of touch on extrastriate visual areas take current gaze direction into account. *Neuron* 34:647–658
- Macaluso E, Frith CD, Driver J (2002b) Supramodal effects of covert spatial orienting triggered by visual or tactile events. *J Cogn Neurosci* 14(3):215–229
- Maravita A, Spence C, Kennett S, Driver J (2002) Tool-use changes multimodal spatial interactions between vision and touch in normal humans. *Cognition* 83(2):B25–B34
- Maravita A, Spence C, Driver J (2003) Multisensory integration and the body schema: close to hand and within reach. *Curr Biol* 13:531–539
- Marini F, Chelazzi L, Maravita A (2013) The costly filtering of potential distraction: evidence for a supramodal mechanism. *J Exp Psychol Gen* 142(3):906–922
- Marini F, Tagliabue CF, Sposito AV, Hernandez-Arieta A, Brugger P, Estévez N, Maravita A (2014) Crossmodal representation of a functional robotic hand arises after extensive training in healthy participants. *Neuropsychologia* 53:178–186
- Mordkoff JT, Egeth HE (1993) Response time and accuracy revisited: converging support for the interactive race model. *J Exp Psychol Hum Percept Perform* 19(5):981
- Murray MM, Wallace MT (eds) (2011) The neural bases of multisensory processes. CRC Press, Boca Raton
- Pavani F, Castiello U (2004) Binding personal and extrapersonal space through body shadows. *Nat Neurosci* 7(1):14–16
- Pavani F, Spence C, Driver J (2000) Visual capture of touch: out-of-the-body experiences with rubber gloves. *Psychol Sci* 11:353–359
- R Core Team (2016) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project>

- Ratcliff R (1993) Methods for dealing with reaction times outliers. *Psychol Bull* 114:510–532
- Sengül A, Rognini G, van Elk M, Aspell JE, Bleuler H, Blanke O (2013) Force feedback facilitates multisensory integration during robotic tool use. *Exp Brain Res* 227(4):497–507
- Shore DI, Barnes ME, Spence C (2006) Temporal aspects of the visuotactile congruency effect. *Neurosci Lett* 392:96–100
- Simon JR (1969) Reactions toward the source of stimulation. *J Exp Psychol* 81(1):174
- Spence C, Kingstone A, Shore DI, Gazzaniga MS (2001) Representation of visuotactile space in the split brain. *Psychol Sci* 12:90–93
- Spence C, Pavani F, Driver J (2004a) Spatial constraints on visual-tactile crossmodal distractor congruency effects. *Cogn, Affect Behav Neurosci* 4:148–169
- Spence C, Pavani F, Maravita A, Holmes N (2004b) Multisensory contributions to the 3-D representation of visuotactile peripersonal space in humans: evidence from the crossmodal congruency task. *J Physiol Paris* 98(1–3):171–189
- Spence C, Pavani F, Maravita A, Holmes N (2007) Multi-sensory interactions. In: Lin MC, Otaduy M (eds) *Haptic rendering: foundations, algorithms, and applications*. AK Peters, Wellesley
- Stein BE, Meredith MA (1993) *The merging of the senses*. The MIT Press, Cambridge
- Stroop JR (1935) Studies of interference in serial verbal reactions. *J Exp Psychol* 18(6):643
- Talsma D, Senkowski D, Soto-Faraco S, Woldorff MG (2010) The multifaceted interplay between attention and multisensory integration. *Trends Cogn Sci* 14(9):400–410
- Teasdale N, Bard C, LaRue J, Fleury M (1993) On the cognitive penetrability of posture control. *Exp Aging Res* 19(1):1–13
- Van Elk M, Forget J, Blanke O (2013) The effect of limb crossing and limb congruency on multisensory integration in peripersonal space for the upper and lower extremities. *Conscious Cogn* 22(2):545–555
- Vecera SP, Farah MJ (1994) Does visual attention select objects or locations? *J Exp Psychol Gen* 123:146–160
- Vercruyssen M, Simonton K (1994) Effects of posture on mental performance: we think faster on our feet than on our seat. In: Lueder R, Noro K (eds) *Hard facts about soft machines—the ergonomics of seating*. Taylor and Francis, London, pp 119–131
- Wallace MT, Stein BE (1997) Development of multisensory neurons and multisensory integration in cat superior colliculus. *J Neurosci* 17(7):2429–2444
- Walton M, Spence C (2004) Crossmodal congruency and visual capture in a visual elevation discrimination task. *Exp Brain Res* 154:113–120
- Wijnen JG, Ridderinkhof KR (2007) Response inhibition in motor and oculomotor conflict tasks: different mechanisms, different dynamics? *Brain Cogn* 63(3):260–270
- World Medical Organization (1996) Declaration of Helsinki. *Br Med J* 313(7070):1448–1449
- Zimmer U, Macaluso E (2007) Processing of multisensory spatial congruency can be dissociated from working memory and visuo-spatial attention. *Eur J Neurosci* 26:1681–1691
- Zopf R, Savage G, Williams MA (2010) Crossmodal congruency measures of lateral distance effects on the rubber hand illusion. *Neuropsychologia* 48(3):713–725
- Zopf R, Savage G, Williams MA (2013) The crossmodal congruency task as a means to obtain an objective behavioral measure in the rubber hand illusion paradigm. *J Vis Exp* 26:77