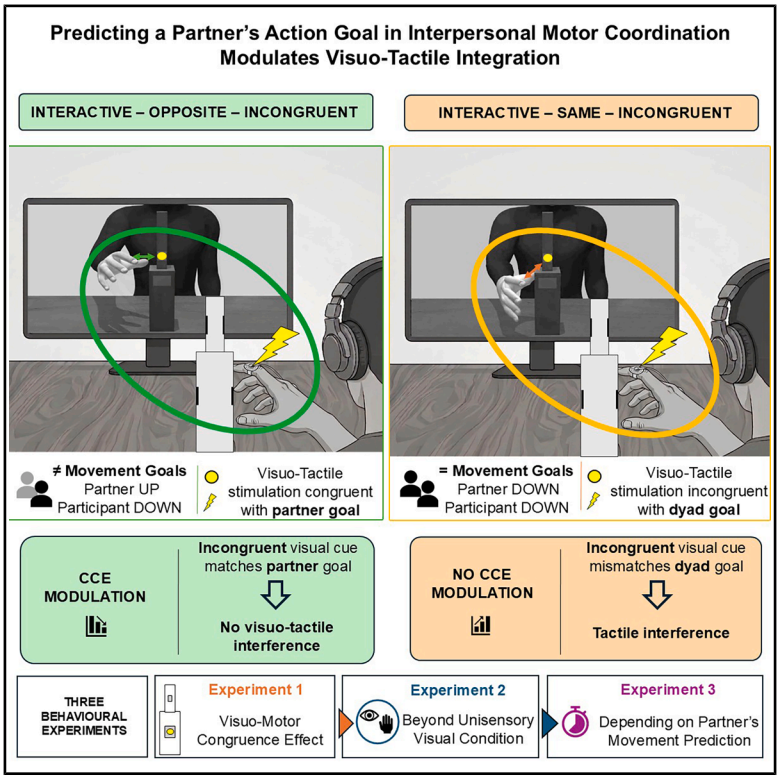
iScience

Article

Interpersonal motor coordination Modulates Visuo-Tactile integration

Graphical abstract



Highlights

- Interpersonal motor interactions modulate cross-modal congruency effects (CCEs)
- Same/opposite interpersonal motor interactions induce opposite modulations of CCE
- Interpersonal motor interactions affect CCE only when predicting partner's action
- Interpersonal motor interactions also modulate unisensory visual processing

Authors

Giovanna Cuomo, Antonio Pio Afragola, Ugo Giulio Pesci, Vanessa Era, Matteo Candidi

Correspondence

giovanna.cuomo@uniroma1.it (G.C.), matteo.candidi@uniroma1.it (M.C.)

In brief

Neuroscience; Cognitive neuroscience

Cuomo et al., 2026, iScience *29*, 116950
August 21, 2026
https://doi.org/10.1016/j.isci.2026.116950

CellPress

iScience

CellPress
OPEN ACCESS

Article

Interpersonal motor coordination Modulates Visuo-Tactile integration

Giovanna Cuomo,[1,2,*] Antonio Pio Afragola,[1] Ugo Giulio Pesci,[1,2] Vanessa Era,[1,2] and Matteo Candidi[1,2,3,*]
[1]Department of Psychology, Sapienza University of Rome, Rome, Italy
[2]IRCCS Fondazione Santa Lucia, Rome, Italy
[3]Lead contact
*Correspondence: giovanna.cuomo@uniroma1.it (G.C.), matteo.candidi@uniroma1.it (M.C.)
https://doi.org/10.1016/j.isci.2026.116950

SUMMARY

Visuo-tactile cross-modal congruence effects (CCEs) refer to impaired tactile processing when task-irrelevant visual cues are presented close to the body but in a spatially incongruent location. This interference extends to visual cues appearing far from the body and on distant objects when individuals are reaching toward those objects or observing someone else grasping them. These effects highlight the impact of motor processes on cross-modal integration. However, it remains unclear whether real-time interpersonal interactions, which require continuous prediction of others' actions, influence the spatial boundaries of CCE. Across three behavioral experiments, we found that interpersonal motor interactions modulate cross-modal integration. Specifically, action-related visuo-tactile CCE emerged only when the location of the visual cue was incongruent with the predicted movement of an interaction partner. These results suggest that interpersonal sensorimotor predictions may shape basic spatial mappings of cross-modal interference, indicating that intersensory processing is dynamically influenced by social interaction demands.

INTRODUCTION

Events that happen close to the body are extremely important because their efficient perception and representation guide goal-directed[1–4] and defensive behaviors.[5] The representation of events and objects in the region of space immediately surrounding a body part,[6] a portion of space called peri-personal (PPS),[7–10] peri-face,[9] peri-hand[10] and peri-trunk space,[6] is supported by a number of multisensory mechanisms that generate intersensory interference and facilitation effects.[7,8,11,12] A classic example of these intersensory effects is the cross-modal congruence effect (CCE).[12,13] This effect is observed when spatially incongruent stimuli presented in a different sensory modality slow down response times and/or increase errors when participants detect a visual[14,15] or auditory cue[16] appearing near a body part or approaching it,[16–19] but located in a different spatial position.

There is extensive evidence that multimodal fronto-parietal neurons and neurons[20–22] in the putamen[1,23,24] integrate visuo-tactile stimuli close to the body and underpin these effects. These neural systems support the behavioral evidence of somatotopically organized peri-hand,[10] peri-trunk[6], and peri-face[9] spaces. Not only are these portions of space somatotopically organized, but their representation is plastically influenced by the way individuals interact with the environment. Indeed, training to use a rack to retrieve food[25] or perform actions,[3,4] may plastically expand the visual receptive field of visuo-tactile neurons and concurrently expand the PPS to include the object of the interaction.[26]

More recently, in line with the plastic nature of the extent of the PPS, it has been shown that visual cues appearing outside the PPS can still facilitate tactile processing, for example when the visual cue appears far from the body, but in correspondence with the location of an intended action, both during its execution and planning.[27–30] This CCE effect associated with movement planning and execution has also been described when participants observe other people grasping objects that belong to the observed person.[31] While the mere presence of another individual in front of us is able to "expand"[32] and "contract"[33] the spatial extent of the PPS, by modulating perception of tactile stimuli over our body,[34,35] theoretical perspectives have suggested that interacting with others might create a "Shared Action Space"[36] (also see Fanghella et al.[37] for a review), interpreted as a space where sensorimotor representations are modulated by the presence and the movements of another individual. One possibility explored in the present study is that multisensory events occurring in the space of an interpersonal interaction are processed and integrated according to different spatial and somatotopic compatibility rules. These rules may differ from those at play when individuals are static, acting alone, or moving asynchronously to another individual. This is in line with the notion that, when performing movements with a partner, we need to be extremely accurate in predicting the sensory (i.e., visual, auditory, and tactile) consequences of their movements and integrate them with the sensory consequences of our own movements.

Across three behavioral experiments, participants performed either a tactile localization (Experiment 1) or a tactile detection

task (Experiments 2 and 3) while engaged in an interpersonal motor task. We tested the hypothesis that, during online coordination with a partner—specifically when performing same or opposite reaching-and-grasping movements toward an object[38–45] (see Figure 1)—the visuo-tactile CCE observed during grasping actions[28–31] is remapped according to the partner's movement. Here, CCE refers to the interference in tactile processing induced by visual cues appearing at an incongruent location from the reaching and grasping target. Specifically, participants performed reaching-and-grasping movements toward a bottle-shaped object, using either a precision or a power grip (targeting the upper or lower part of the object), while simultaneously synchronizing and coordinating with a virtual partner by performing the same or opposite movements (Figure 1A). During this interpersonal motor task, a tactile stimulus was delivered to the finger that was about to contact the object, while a visual cue appeared on the partner's target object at a location that could be either congruent or incongruent with the target location of the participant's or the partner's movement (Figure 1B). We tested whether the interference of visual cues on tactile processing is modulated by the type of interpersonal motor coordination and by the movement target location of the participant and the partner. Four experimental factors were manipulated: (1) interaction type, in which participants either needed to predict the partner's upcoming movement to perform the same or opposite grasping action (interactive condition) or followed a spatial instruction indicating where to grasp (cued condition); (2) movement type, in which the movements performed by the participant resulted in either the same or opposite movements compared with those of the virtual partner (e.g., both partners grasp with a precision grip at the top of the bottles, or one with a precision and the other with a power grip); (3) VisuoMotor congruence, related to the visual cue appearing at a location that was either spatially congruent or incongruent with the participant's end-movement target location (VM-congruent vs. VM-incongruent); (4) only in Experiment 1 we included a factor named somatotopic congruence, in which the visual cue could appear on either the front or back of the partner's bottle so that it was either somatically congruent or incongruent with the participant's finger that received the tactile stimulus (i.e., thumb or index, respectively), depending on the end grasp location on the bottle (somatotopic-congruent vs. somatotopic-incongruent).

The three experiments investigated whether and how online interpersonal motor interactions shape visuo-tactile CCE, defined as the difference in reaction times to tactile stimuli between VisuoMotor congruent and incongruent trials. In congruent trials, the visual cue appears at the spatial location reached by the finger receiving the tactile stimulus, whereas in incongruent trials it appears at a different location. While this manipulation was constant in all three experiments, each experiment also implemented a specific manipulation to test the somatotopy of the CCEs (Experiment 1), their cross-modal nature (Experiment 2) and their dependency on interpersonal predictions (Experiment 3).

Experiment 1 investigated whether the classical CCE is modulated when performing interpersonal motor interactions. Participants were asked to perform a tactile localization task (i.e., whether they perceived a tactile stimulus on their index or thumb) while executing a reach-to-grasp movement toward a bottle in front of them in synchrony with a virtual partner. Cross-modal (in)congruence was implemented in the experimental set-up in two ways. First, by presenting task-irrelevant visual cues on either the upper or lower part of the object that will be grasped by the partner (thus in the same or opposite spatial location of one's own movement), generating a VisuoMotor congruence factor (Figure 2A). Second, by presenting these task-irrelevant visual cues either on the front or back of the partner's bottle, thus in spatially congruent (thumb and index, respectively) or incongruent (index and thumb, respectively) locations with the finger on which participants receive the tactile stimulus (i.e., Somatotopic congruence; Figure 2B). In relation to the first spatial incongruence (i.e., VisuoMotor congruence), we hypothesized that cross-modal (in)congruence would be evident only during VisuoMotor incongruent trials when engaged in same interactions in a realistic interaction (interactive condition) compared to when performing opposite interaction in the same condition or during the control condition (Cued condition). Specifically, during interactive same reach-to-grasp movements, we expect the classical CCE: Impaired tactile localization when a task-irrelevant visual cue appears at a location incongruent with the movement's target.[28–31] In contrast, for opposite reach-to-grasp interactions, we hypothesize that tactile detection will no longer be disrupted by spatially incongruent visual cues, since these cues are now congruent with the partner's movement. This pattern would be consistent with an interpersonal remapping of visuo-tactile integration. We did not expect such modulation in a cued condition, where the specific partner's movement was not relevant for the participant's action goal. The second type of spatial incongruence (i.e., Somatotopic congruence) was meant to test whether the visual cues appearing on the object of the partner would be mapped as if the object was grasped by the participant. Thus, a cue on the front of the bottle is congruent with the thumb of the participants, and a visual cue on the back of the bottle is congruent with their index, or in relation to the movement of the partner, implying the opposite congruence logic, i.e., front-index and back-thumb.

Experiment 2 aimed at examining whether the CCE modulations observed in Experiment 1 reflect a visuo-tactile, motor-dependent, cross-modal effect and extend beyond a unisensory (visual) motor-dependent effect. In different experimental blocks, participants performed a tactile detection task either during a visuo-tactile stimulation or a visual stimulation (i.e., visuo-tactile vs. visual condition) while observing the partner's movement. As in Experiment 1, in the visuo-tactile condition, the tactile stimulus was delivered along with a task-irrelevant visual cue (appearing on the upper or lower part of the bottle), which was either congruent or incongruent with the direction of their own and/or the partner's movement (i.e., VisuoMotor congruence). In the visual condition, participants were asked to detect a visual cue appearing at the same locations on the upper or lower part of the partner bottle without receiving any tactile stimulus. Unlike Experiment 1, no somatic mapping between task-irrelevant visual cues and

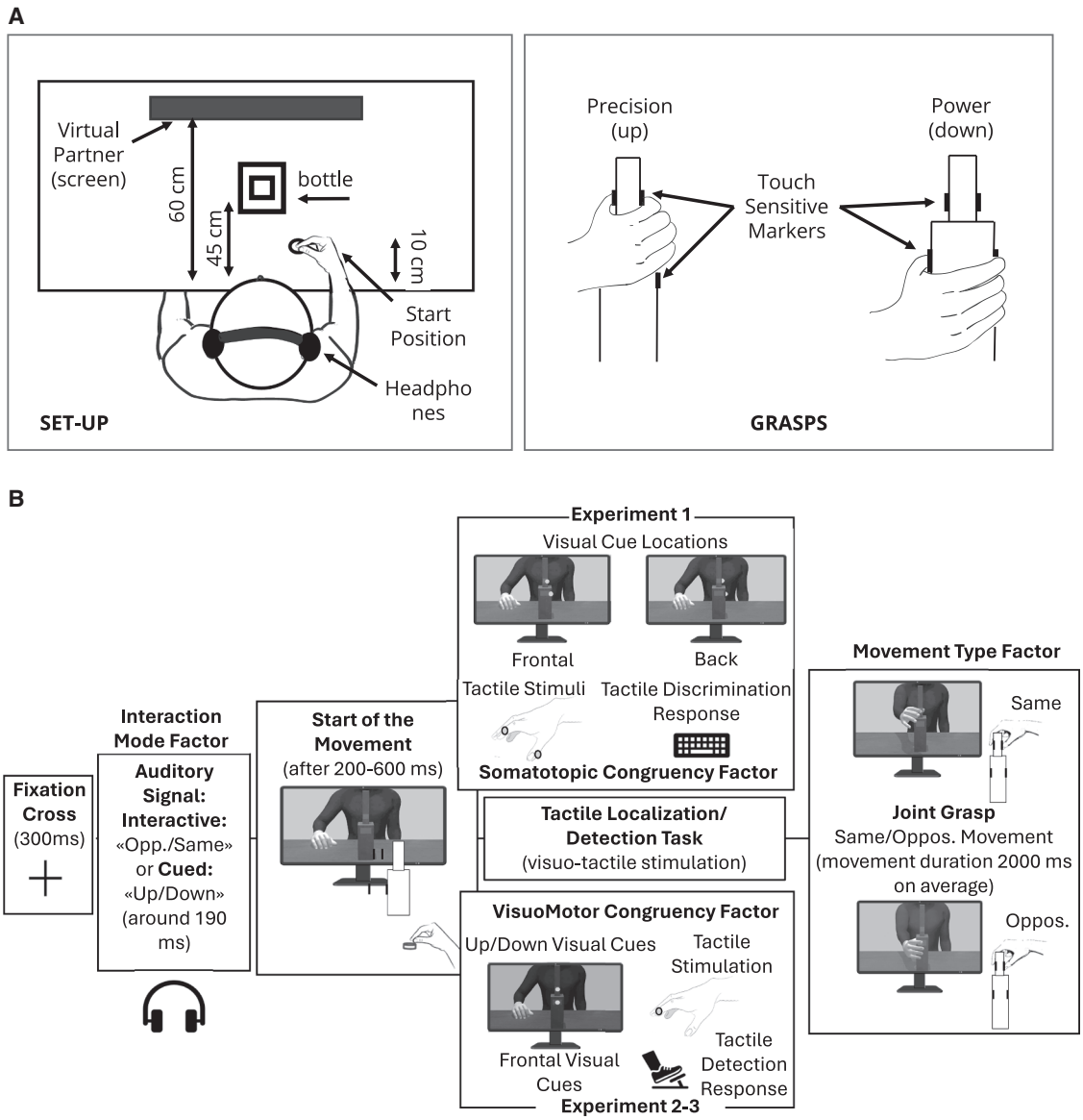


Figure 1. General experimental set-up

(A) Experimental set-up. Participants sat at a table facing a screen displaying a virtual partner. They were instructed to synchronize their grasping action on a bottle—performing either a precision grip (up) or a power grasp (down)—with the virtual partner.

(B) Trial timeline. In all three experiments, participants started by placing their thumb and index fingers on a button. They were either instructed to predict the partner's movement (interactive condition) and perform the same or opposite action compared to the partner's, or directed to a specific location (cued condition), but were always told to synchronize their grasp time with that of the partner. While moving with the partner, they performed a tactile localization/detection task while also a visual cue flashed on the upper or lower part of the partner's bottle, corresponding or not to their grasping goal (VisuoMotor congruence) and on the front or back of the bottle (Somatotopic congruence; only Experiment 1), corresponding or not to their finger contact position on the bottle. The movement ended in a joint grasp, defined as same or opposite relative to the partner (movement type).

the location of tactile finger stimuli was present. If the social modulation of CCE described in Experiment 1 reflects cross-modal processing rather than visual attention alone, we expected the CCE to appear in both the visual and the visuo-tactile conditions.

Experiment 3 aimed to test the interactive nature of the visuo-tactile CCE observed in the previous two studies. Specifically, we manipulated the timing of the visuo-tactile stimulation as a function of the phase of the movement of the partner. Visuo-tactile stimuli were thus delivered when the phase of the partner's movement allowed (late visuo-tactile timing) or did not allow (early visuo-tactile timing) participants to predict the location of the partner's movement. We hypothesized that CCE modulation would not emerge at the initial phase of the movement when upper or lower reaching movements cannot be discriminated (early visuo-tactile timing), and only emerge

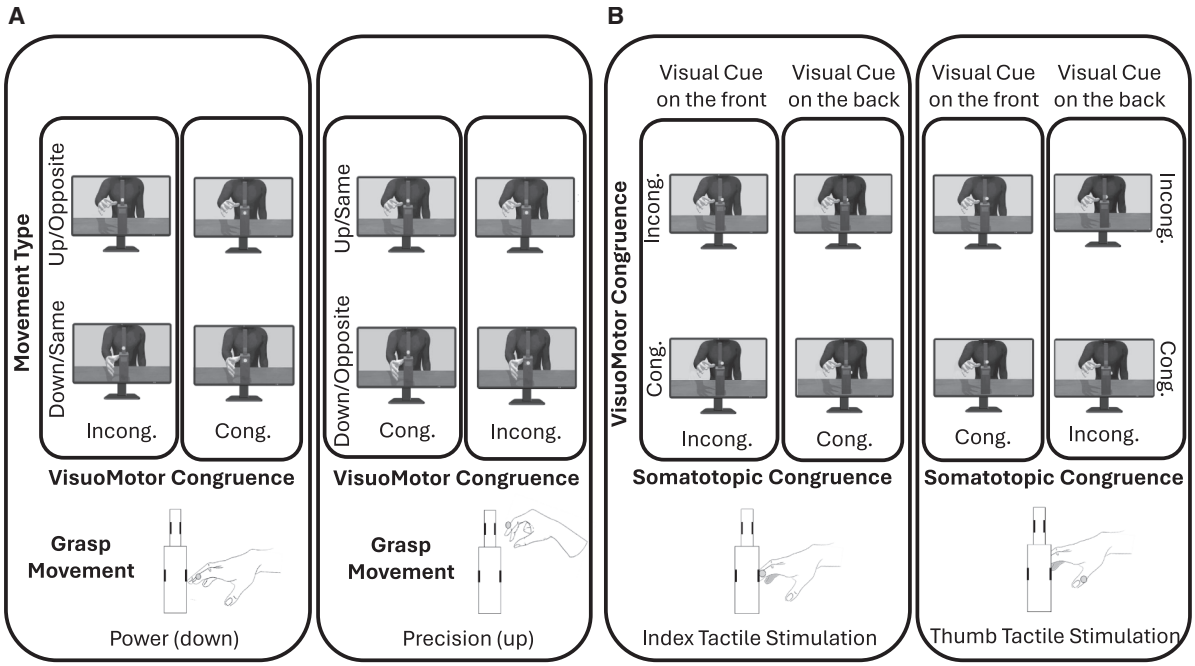


Figure 2. Illustration of VisuoMotor (VM) and Somatotopic congruency factors

(A) VisuoMotor (VM) congruence factor. Participants coordinated their reach-to-grasp action (power—left panel; precision—right panel) with those of a virtual partner, resulting in performing same or opposite movement relative to the partner's ones. The visual cues appeared in congruent or incongruent positions relative to participants' movement goal (on the upper/lower portion of the bottle). This condition was present across all three experiments and regarded the object vertical axis.

(B) Somatotopic congruence factor. Participants received tactile stimuli on their index (left panel) or thumb (right panel) while visual cues appeared on the virtual partner's bottle in positions that were either somatotopically congruent or incongruent with their own end-finger-position on the bottle (front/back). This factor was only present in Experiment 1. For simplicity the image depicts an example of power grasping movement combinations only.

at the late timing (late visuo-tactile timing). This would support the idea that online action prediction of the partner's movement is necessary to elicit the interpersonal remapping of multisensory processes.

Collectively, testing these hypotheses provides new insights into the flexibility of sensory mechanisms at play during interpersonal interactions.

RESULTS

For each Experiment, we report in the following sections the main effect of VisuoMotor congruence, indexing the emerge of the CCE together with the highest-order interaction involving interaction type, movement type and VisuoMotor congruence, as this reflects the primary effect of interest. Additionally, significant effects that are relevant for the interpretation of the results are also reported. All other significant results are provided in detail in the Tables in the supplemental information.

Reaction times (RTs) were collected for the tactile localization/detection task, while grasping asynchrony (GA) was collected for the interpersonal motor task and is reported as supplemental information. Error rate was considered as the incorrect or missed identification of the tactile stimulus (localization in Experiment 1; detection in Experiments 2 and 3), limited to trials with correct joint grasp execution (range 0–1).

Experiment 1—Interpersonal motor interactions impact the CCE

In this experiment, we explored whether visuo-tactile CCE observed during grasping actions is modulated by interpersonal motor coordination. Specifically, we examined whether tactile localization of a stimulus (presented on the index or the thumb) during partner's movement prediction was influenced by: (1) the spatial relationship between the visual cue appearance and either the participant's or the partner's movement goal (VisuoMotor congruence factor; Figure 2A) as a function of the movement type factor, and (2) by the somatotopic correspondence between the visual cue on the partner's object and the stimulated finger of the participant (i.e., the location that corresponded to where the subject index or thumb would touch the partner's object; Somatotopic congruence factor; Figure 2B).

Reaction times results of the tactile localization task

We conducted a $2 \times 2 \times 2 \times 2$ within-subject repeated measures analysis of variance (ANOVA) on participants' reaction times with the following factors: Interaction type (cued/interactive), Movement type (opposite/same), VisuoMotor congruence (VM-congruent/VM-incongruent; i.e., the congruence of the visual cue was defined in terms of its spatial relation to the location of the up or down reaching movement of the participant) and Somatotopic congruence (S-congruent/S-incongruent, i.e., the congruence of the visual cue was defined in terms of its relation

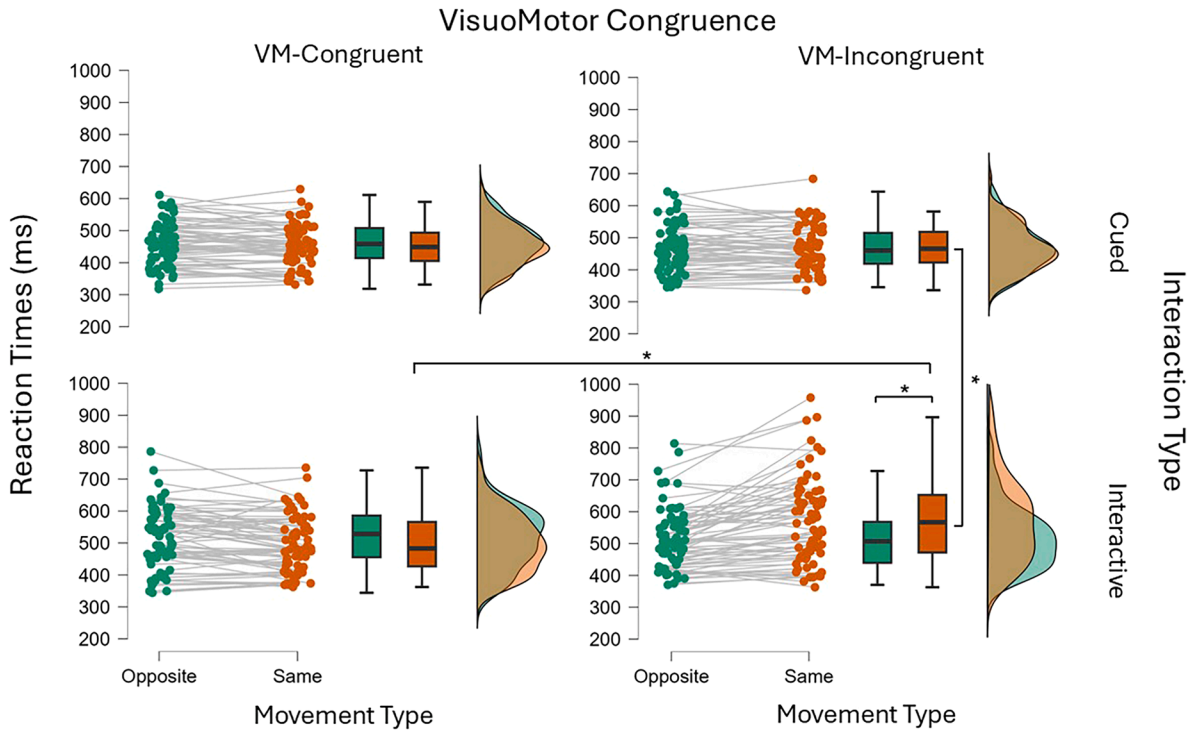


Figure 3. Significant interaction between Interaction type (cued/interactive) * Movement type (same/opposite) * VisuoMotor congruence (VM-congruent/VM-incongruent) on RTs (ms) (N = 34, F(1, 33) = 12.88, p = 0.001, ηp^2 = 0.28)

Asterisks indicate significant differences among conditions at $p < 0.05$ Bonferroni-corrected in the ANOVA. Boxplots indicate the median and interquartile range; density plots show individual variability; whiskers show 95% confidence intervals.

to the location of the tactile stimulus on the participant's hand, i.e., index or thumb) (all effects are reported in Table S1).

The analysis showed a significant main effect of VisuoMotor congruence ($F(1,33) = 31.08$, $p < 0.001$, $\eta_p^2 = 0.49$) with longer reaction times for VisuoMotor incongruent compared to congruent trials testifying the presence of a CCE which was specified by the three-way interaction between Interaction mode (cued/interactive) × Movement type (opposite/same) × VisuoMotor congruence (VM-congruent/VM-incongruent).

The 2 × 2 × 2 Interaction mode (cued/interactive) × Movement type (opposite/same) × VisuoMotor congruence (VM-congruent/VM-incongruent) interaction resulted significant ($F(1,33) = 12.88$, $p = 0.001$, $\eta_p^2 = 0.28$) confirming the presence of CCE in our experimental set-up only in some conditions of the interaction (Figure 3). Bonferroni-corrected post-hoc analyses (Table S2) revealed that the slowing of RTs during VM-Incongruent trials was statistically significant only in the interactive condition when performing same movements with the partner (M = 574.12 ms; SD = 131.94 ms) compared to VM-congruent ones (M = 499.05 ms; SD = 86.24 ms) ($p < 0.001$). This pattern was consistent with the notion that a visual cue appearing at a location different from that of the individual's (and partner) action goals impairs the processing of a tactile stimulus. Crucially, however, this interferent effect (CCE) was not found during the execution of opposite movements in the interactive condition ($p > 0.9$). This suggests that during interpersonal online interactions, the interference generated by spatially incongruent visual cues on the ability to detect a tactile stimulus observed in same movements was not present when performing opposite actions compared to the partner (Figure 3). This interpretation was further confirmed by the presence of significantly slower RTs in interactive VM-incongruent trials during same movements (M = 574.12 ms; SD = 131.94 ms) compared to opposite ones (M = 516.57 ms; SD = 95.87 ms) ($p < 0.001$) where the visual cue would still be incongruent with the participant's motor aim, but congruent with the action goal of the partner (Figure 3).

Moreover, these patterns of results were not present in the cued condition, where tactile RTs of same VM-incongruent trials (M = 468.27 ms, SD = 67.21 ms) did not significantly differ either from VM-congruent ones (M = 450.92 ms, SD = 63.91 ms; $p > 0.9$), nor from opposite VM-incongruent ones (M = 467.95 ms, SD = 68.56 ms; $p > 0.9$). This pattern suggests that the type of interaction with the partner during joint grasping was fundamental in shaping the CCE observed in the Interactive conditions. Furthermore, RTs of the interactive same VM-incongruent condition (M = 574.12 ms; SD = 131.94 ms) were significantly slower compared to those of the cued one (M = 468.27 ms, SD = 67.21 ms; $p < 0.001$), indicating that the interference effect triggered by the VisuoMotor incongruence emerged specifically during interpersonal coordination, when participants had to continuously monitor and anticipate the partner's actions.

The interaction between Somatotopic congruence (S-congruent/S-incongruent) and Movement type factors (same/opposite) was significant ($F(1, 33) = 13.19$, $p = 0.0009$, $\eta_p^2 = 0.29$) (Figure 4).

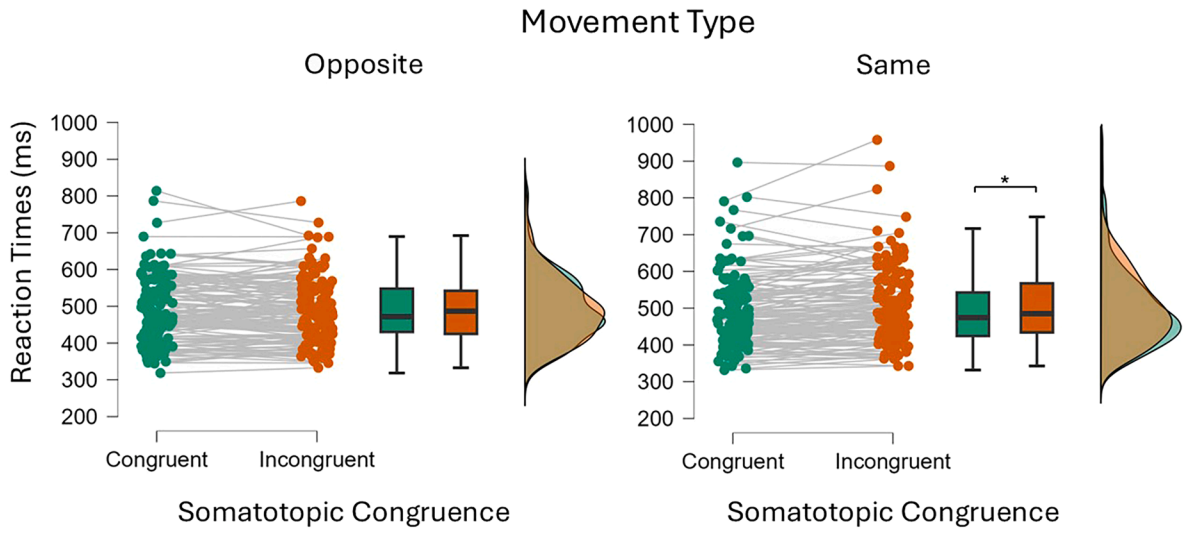


Figure 4. Significant interaction between Movement type (same/opposite) * Somatotopic congruence (S-congruent/S-incongruent) on RTs (ms) (N = 34, F(1, 33) = 13.19, $p < 0.001$, $\eta p^2 = 0.29$)

Asterisks indicate significant differences among conditions at $p < 0.05$, Bonferroni-corrected in the ANOVA. Boxplots indicate the median and interquartile range; density plots show individual variability; whiskers represent the 95% confidence interval.

Post-hoc analysis (Table S3) indicated that only during the execution of same movements, somatotopic incongruent trials (M = 505.51 ms; SD = 81.98 ms) resulted in an interference effect compared to congruent ones (M = 490.67 ms; SD = 78.54 ms; p = 0.0002) (i.e., visual cue appearing on the partner's object where the participant's finger (e.g., index) would make contact interfered with localization of tactile stimuli on the other finger (thumb)), regardless of whether participants performed interactive or cued trials and regardless of whether visual cues appeared in a location that was S-congruent or S-incongruent with the movements of the participant. This pattern of results suggests that during interpersonal interactions that entail performing the same actions as a partner, individuals tend to code the spatial location of visual cues according to a somatotopic frame of reference during same movements (see discussion Experiment 1).

All other significant interactions details can be consulted in the Tables S1–S3.

Error rate results of the tactile localization task

The 2 × 2 × 2 × 2 aligned rank transform (ART) repeated measure ANOVA on the error rate of the tactile stimulation localization, with factors Interaction type (cued/interactive) × Movement type (opposite/same) × VisuoMotor congruence (VM-congruent/VM-incongruent) and Somatotopic congruence (S-congruent/S-incongruent) revealed a significant three-way interaction between interaction type, movement type, and VisuoMotor congruence (F(1, 495) = 16.20, $p < 0.001$, $\eta_p^2 = 0.03$) (Table S4). To further explore this three-way interaction, Bonferroni-corrected post hoc comparisons were conducted (Table S5).

Error rate was significantly higher in the interactive same incongruent condition (M = 0.28, SD = 0.15) than in cued same incongruent (M = 0.17, SD = 0.12; $p < 0.001$). Within the interactive type, error rate was significantly higher in same incongruent (M = 0.28, SD = 0.15) compared to congruent trials (M = 0.15, SD = 0.12; $p < 0.001$), confirming a specific cost for cross-modal incongruence during mutual interaction already observed in RTs, that is not found for opposite movements ($p > 0.99$) (Figure 5).

In summary, the results of Experiment 1 suggest that participants performed significantly worse in the interactive condition, when there were incongruencies between visual-cue and motor information, especially during same movements, but not during opposite ones. These patterns were consistent with the hypothesis that dyadic coordination modulates the integration of multisensory and motor information, and that shared action goals might influence the processing of task-irrelevant stimuli presented in the partner's action space.

This first experiment aimed to investigate whether, during synchronous interpersonal motor interactions, task-irrelevant visual cues modulated tactile localization on the acting hand as a function of spatial congruence. Specifically, visual cues could appear at locations that are congruent or incongruent with either (1) the spatial target of one's own movement, (2) the partner's movement, or (3) the somatotopic location of a tactile stimulus on the participant's finger. The results highlight that, during an interpersonal interaction, the impact of the location of visual cues on individuals' ability to discriminate the location of a tactile stimulus does no longer depend on its spatial congruence with the target of individual's movement as shown in previous studies,[28–31] but it depends on its congruence with respect to the spatial location of the predicted target movement of the partner. Specifically, an interfering effect of the visual cue that appeared at a spatial location that is incongruent with one's own movement only manifested if it was necessary to predict the partner's actions (interactive type) during same interactions type. In contrast, during opposite movements, the interfering effect of the visual cue was not observed when presented in an incongruent spatial location relative to where the subject needed to grasp, but congruent to the end location of the partner's movement. This would possibly occur because the participant

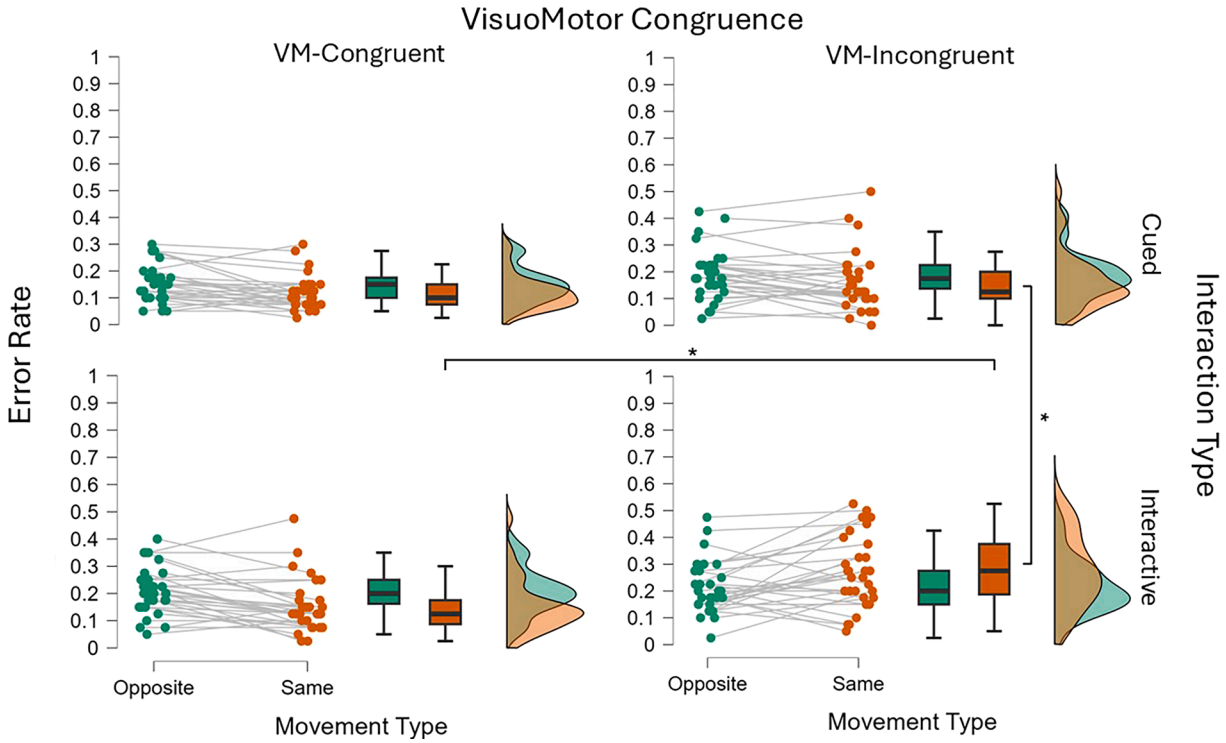


Figure 5. Significant interaction between Interaction type (cued/interactive) * Movement type (same/opposite) * VisuoMotor congruence (VM-congruent/VM-incongruent) on error rate (N = 34, $F(1, 495) = 16.20$, $p < 0.001$, $\eta p^2 = 0.29$)

Asterisks indicate significant differences among conditions at $p < 0.05$, Bonferroni-corrected in the ANOVA. Boxplots indicate the median and interquartile range; density plots show individual variability; whiskers represent the 95% confidence interval.

was predicting the partner's action, whose end-movement position corresponded with the portion of space in which the incongruent visual cue was presented. Importantly, such modulation was absent in the cued condition, where participants performed the same interpersonal synchronous motor task without the need to actively predict the partner's movements while coordinating. In this context, the visual cues appearing on the partner's object did not elicit any interference effect during spatially incongruent trials with respect to one's own movement. This suggests that the visual cue processing differed when no active representation of the partner's action was required. Thus, the interfering role of the cue critically depends on whether the participant is engaged in an interactive context where predicting the partner's movement is necessary. These modulatory effects of visual cue congruence based on the interaction context and shared movement execution were consistently observed across our dependent measures (RTs and error rate, also see supplemental information for GA data), highlighting the robustness of the observed CCE linked to interpersonal prediction mechanisms.

Furthermore, during the execution of same movements (regardless of CCE), we also found an interferent visuo-tactile effect of visual cues appearing at an incongruent somatotopic location with the participant's finger that received the tactile stimulation, i.e., the visual cue appears on the back of the partner's bottle, and the participant gets the tactile stimulus on their thumb. This effect is coherent with previous findings by[31] who showed that while observing a person grasping an object that is perceived as owned by them, visual cues appearing on the object might modulate tactile detection of stimuli delivered on the participant's finger according to somatotopic coherence (i.e., visual cue appearing on the partner's object where a participant's finger (e.g., index) would touch it interfere with the localization of tactile stimuli on the other finger (thumb)). This suggested that the CCE effects found in Patanè et al.[31] for objects that were associated with either oneself or with both oneself and other individuals were extended to the companion's object while interacting. Hence, these results might shed light on the nature of low-level motor encoding of interactions in shaping intersensory effects.

Experiment 2—Interpersonal motor interaction affect unisensory processing and CCE

To determine whether the emergence of a CCE during same movement execution in the interactive condition found in Experiment 1 reflected a cross-modal phenomenon rather than a purely visual effect (e.g., facilitatory effect of the movement of the partner on the detection of the task-irrelevant visual cue), we conducted a second experiment including a unisensory visual condition to compare to the Visuo-Tactile one. Thus, in Experiment 2, participants were asked to perform two different sensory detection tasks and either respond as fast as possible to a visual cue appearing on the upper or lower part of the partner's object in absence of any tactile stimulus (unisensory visual condition), or to a tactile stimulus on their index finger while

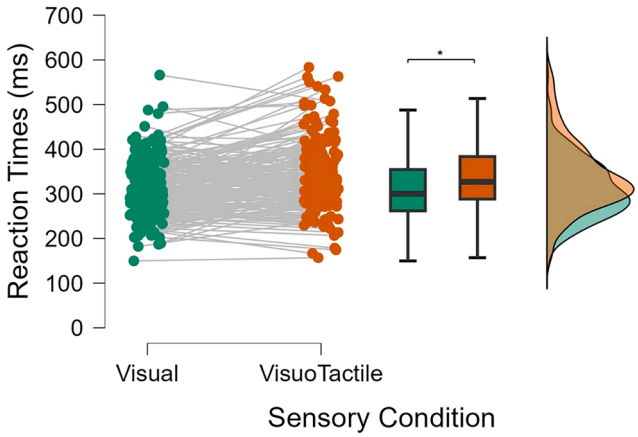


Figure 6. Main effect of the Sensory condition (visual/visuo-tactile) on RTs (in ms) (N = 34, F(1, 23) = 9.44, p = 0.01, ηp^2 = 0.29)

Asterisks indicate significant differences among conditions at $p < 0.05$, Bonferroni corrected in the ANOVA. Boxplots indicate the median and interquartile range; density plots show individual variability; whiskers represent 95% confidence intervals.

ignoring the co-occurring task-irrelevant visual cue (cross-modal visuo-tactile condition). Critically as in Experiment 1, in both sensory conditions, visual cues were either congruent or incongruent with the direction of the participant's own reaching movement (i.e., VisuoMotor congruence). Unlike Experiment 1, no somatotopic correspondence between the visual cue and the tactile location was present.

We hypothesized that a difference between the unisensory visual and the cross-modal visuo-tactile tasks would strengthen the evidence for an intersensory effect found in Experiment 1, not reducible to enhanced visual saliency or attentional capture by the partner's movement at the task-irrelevant visual cue location. Furthermore, if participants were simply reacting to the visual cue in the visuo-tactile condition regardless of task instructions, we would expect no differences between response times across the visual and visuo-tactile conditions.

Reaction times results of the tactile detection task

The 2 × 2 × 2 × 2 within-subject repeated measures ANOVA on RTs for the tactile task included the following factors: Interaction type (cued/interactive) × Sensory condition (visual/visuo-tactile) × VisuoMotor congruence (VM-congruent/VM-incongruent) × Movement type (opposite/same) (Table S8).

The analysis revealed a main effect of Sensory condition ($F(1,23) = 9.44$, $p = 0.01$, $\eta_p^2 = 0.29$), indicating significantly faster RTs in the visual condition (M = 324.32 ms, SD = 34.80 ms) compared to the visuo-tactile one (M = 359.70 ms, SD = 22.23 ms). This main effect (Figure 6) is relevant as it supports the idea that when performing the visuo-tactile condition, participants were paying attention to the tactile stimulus rather than ignoring it and responding to the visual one.

The ANOVA also showed a significant main effect of VisuoMotor congruence ($F(1,23) = 23.97$, $p = 0.0001$, $\eta_p^2 = 0.51$) testifying to the emergence of a CCE in our experiment. This main effect was further specified by the significant third-level interaction between Interaction type (cued/interactive), Movement type (opposite/Same) and VisuoMotor congruence (VM-congruent/VM-incongruent) ($F(1,23) = 5.87$, $p = 0.02$, $\eta_p^2 = 0.20$) (Figure 7). Post-hoc tests (Table S9) confirmed the same pattern of results found in Experiment 1.

These results, thus, replicated the findings of Experiment 1 and support the view that the CCE in this task reflects intersensory interference and cannot be explained by a purely visual effect.

All other significant interaction details can be consulted in the Tables S8 and S9.

Error rate results of the tactile detection task

The 2 × 2 × 2 × 2 ART ANOVA on the error rate of the tactile detection task, including Interaction type (cued/interactive) × Movement type (opposite/same) × VisuoMotor congruence (VM-congruent/VM-incongruent) × Sensory condition (visual/visuo-tactile), revealed a significant main effect of Interaction type ($F(1, 345) = 5.61$, $p = 0.018$, $\eta_p^2 = 0.016$) with higher error rate in the interactive condition (M = 0.165, SD = 0.12) compared to the cued one (M = 0.13, SD = 0.12). No other effects resulted in non-significant results (all ps > 0.06) (Table S10).

The second experiment was designed to investigate whether the CCE, observed in Experiment 1, was a cross-modal phenomenon and was not only explained by a visual attentional facilitatory effect of the movement of the partner on the detection of the task-irrelevant visual cue. The results showed that a spatial incongruence between the participant's movement target and the visual cue on the partner's object impaired performance in detecting tactile stimuli (slower RTs, CCE). The visual cue's incongruent position on the partner's object impaired the participant's tactile response only when there is a need to predict the movement of the partner (interactive mode) during same interactions (same condition) compared to complementary interactions (opposite condition). The main effect of sensory condition is consistent with the interpretation that participants were actually differentiating between the visuo-tactile and the visual task. The visual cue had an effect on tactile processing that cannot be attributed to a visual capture of attention alone, or participants only relying on the visual cue and ignoring the tactile stimulus.

The results pattern suggested that participants' capacity to detect tactile stimuli during interpersonal interactions was influenced by the spatial congruence of visual cues to both their own and the partner's movements. This implied that predicting the partner's movements primed specific spatial locations, potentially heightening visual attention toward those areas that are relevant for the interaction. Overall, these findings suggested that CCE was remapped by the rules of interpersonal motor interactions, replicating the results of Experiment 1 and confirming that these effects similarly impact unisensory and cross-modal visuo-tactile processing.

Experiment 3—Predictions affect CCE during interpersonal motor interaction

In the previous experiments, participants received visuo-tactile stimuli at a single time point during the partner's movement. At this time point (e.g., stage of the movement) participants were able to predict the partner's grasping end-location, as demonstrated in a pilot study (see STAR Methods section and Figure S1). In Experiment 3, we aimed to study whether the

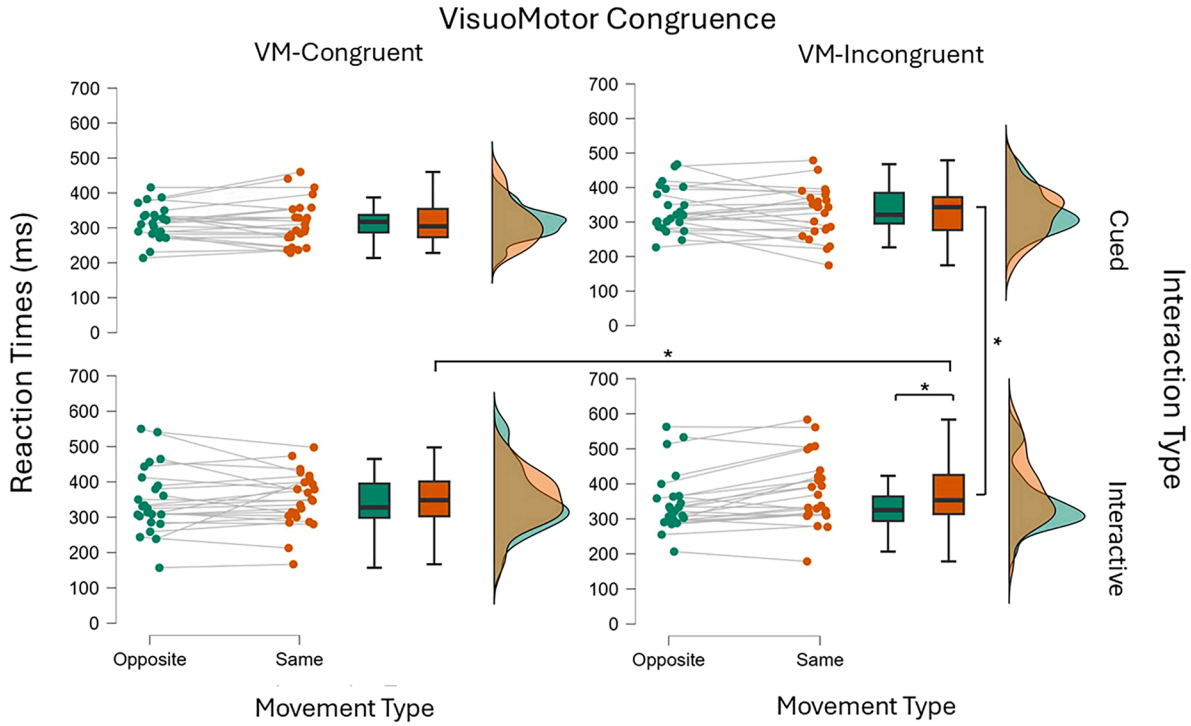


Figure 7. Interaction type (cued/interactive) * Movement type (same/opposite) * VisuoMotor congruence (VM-incongruent/VM-congruent) on RTs (ms) (N = 24, F(1, 23) = 5.87, p = 0.02, ηp^2 = 0.20)

Asterisks indicate significant differences among conditions at $p < 0.05$, Bonferroni corrected in the ANOVA. Boxplots indicate the median and interquartile range; density plots show individual variability; whiskers show 95% confidence intervals.

visuo-tactile CCE would be modulated as a function of the fact that the tactile stimulus was delivered at timings when target location of the partner movement was, or was not, possible to predict by observing the movement of the partner. We thus delivered the visuo-tactile stimuli at two different time points during the movement of the partner. We introduced an early timing of visuo-tactile stimuli delivery that occurred at a time point at which the two different partners' movements (i.e., precision or power grasp) could not be discriminated and predicted, and maintained the second time point as in Experiments 1 and 2. We hypothesized that CCE would emerge during the execution of same movements in the interactive condition only in conditions where the avatar's reach-to-grasp end location was predictable (i.e., when the visuo-tactile stimuli were delivered at the second time point; late tactile timing), and would not occur when the VisuoMotor incongruence is lacking due to the fact that the target of the movement of the partner is not predictable (early tactile timing).

To test these hypotheses, we ran two separate ANOVAs on early and late tactile stimulation timing.

Reaction times results of the tactile detection task at early visuo-tactile stimulation timing

The 2 × 2 × 2 ANOVA on early visuo-tactile stimulation timing showed no CCE as the main effect of VisuoMotor congruence did not turn significant (p = 0.49). Similarly, at early visuo-tactile stimulation timing the three-way Interaction type (cued/interactive) × Movement type (same/opposite), VisuoMotor congruence (VM-congruent/VM-incongruent) was not significant (p = 0.16) (Table S13). We used a Bayesian approach to test the evidence in favor of a difference between interactive VM-incongruent same condition compared to the opposite one in generating slower RTs. The analysis supported the absence of evidence in favor of the alternative hypothesis (B10 = 0.45).

Reaction times results of the tactile detection task at late visuo-tactile stimulation timing

The 2 × 2 × 2 ANOVA on late visuo-tactile stimulation timing instead showed that the three-way Interaction type × Movement type × VisuoMotor congruence resulted significant (F(1,17) = 4.83, p = 0.04, η_p^2 = 0.22) (Figure 8). The post-hoc test pattern replicates the findings of Experiments 1 and 2 (Table S13).

Post-hoc analyses (Table S14) showed significantly slower tactile reaction times (p = 0.03) during same interactions in interactive incongruent motor conditions (M = 461.97 ms, SD = 84.63 ms) compared to opposite ones (M = 417.19 ms, SD = 62.94 ms) (Figure 8). This difference was not present in the cued condition ($p > 0.9$), suggesting that the interaction with the partner of the joint grasping was fundamental in shaping CCE observed in the interactive conditions. Performing same movements during the interactive condition tended to generate slower RTs during motor incongruent trials (M = 461.97 ms, SD = 84.63 ms) compared to congruent ones (M = 419.88 ms; SD = 63.11 ms) (p = 0.06). However, RTs of the cued same VM-incongruent condition (M = 396.33 ms, SD = 43.84 ms) were significantly faster compared to the interactive one (M = 461.97 ms; SD = 84.63 ms; $p < 0.001$),

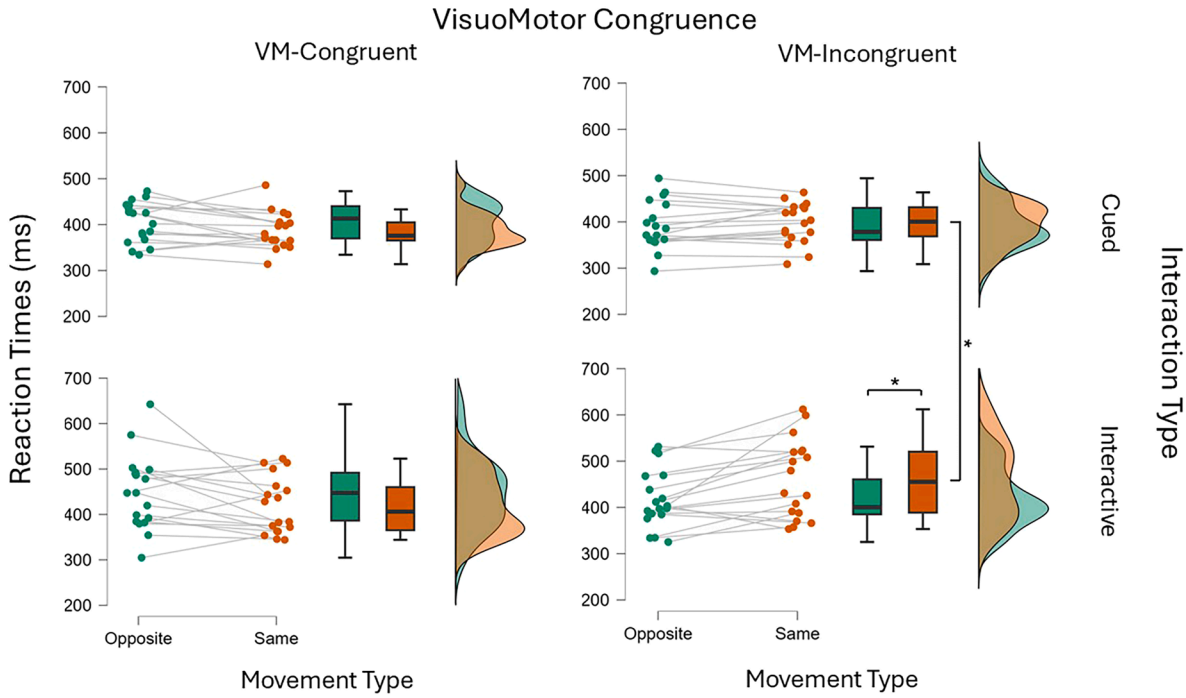


Figure 8. Interaction type (cued/interactive) * Movement type (same/opposite) * VisuoMotor congruence (VM-incongruent/VM-congruent) on RTs (ms) in late visuo-tactile timings (N = 18, F(1, 17) = 4.83, p = 0.04, ηp^2 = 0.22)

Asterisks indicate significant differences among conditions at $p < 0.05$, Bonferroni-corrected in the ANOVA. Boxplots indicate the median and interquartile range; density plots show individual variability; whiskers represent 95% confidence intervals.

indicating that the interference effect triggered by the VM-incongruence emerged specifically during interpersonal coordination (interactive condition), when participants had to continuously monitor and anticipate the partner's actions to perform the same movement. We used a Bayesian approach to test the evidence in favor of an effect of interactive VM-incongruent same conditions compared to opposite ones in generating slower RTs. The analysis supported evidence in favor of the alternative hypothesis (B10 = 7.70).

Overall, these results complemented the findings from Experiments 1 and 2, and supported the modulation of visuo-tactile interference depended on whether information about the partner's upcoming movement target was available.

Error rate of the tactile detection task at early visuo-tactile stimulation timing

The 2 × 2 × 2 Interaction type (cued/interactive) × Movement type (opposite/same) × VisuoMotor congruence (VM-congruent/VM-incongruent) Friedman ANOVA on error rates of the early visuo-tactile stimulation timing revealed no significant difference across conditions (Chi sqr.(N = 18, df = 7) = 6.698 p = 0.461).

Error rate of the tactile detection task at late visuo-tactile stimulation timing

The 2 × 2 × 2 Interaction type (cued/interactive) × Movement type (opposite/same) × VisuoMotor congruence (VM-congruent/VM-incongruent) Friedman ANOVA on error rates of the late visuo-tactile stimulation timing revealed no significant differences across conditions (Chi sqr.(N = 18, df = 8) = 4.432 p = 0.816).

The present results confirmed that the interferent effect of visual cues that were spatially incongruent with one's movement was found only for visuo-tactile stimulation presented at a timing at which participants were able to predict the partner's movement (i.e., late time point).

Relationship between cross-modal congruence effect and behavioral performance at the interaction task

To explore the relationship between performance in the tactile task and the interpersonal motor task, we conducted a correlation between an index of the tactile RTs and the same index on GA of the interpersonal motor task, pooling together data from the three studies. RTs and GA indexes were both obtained by subtracting the individual's mean values in the interactive VisuoMotor incongruent conditions during opposite movement trials from the corresponding values during same movement trials (i.e., RTs/GA index = (interactive VisuoMotor incongruent same movement)—(interactive VisuoMotor incongruent opposite movement)). The index on tactile RTs captured the CCE of visual cues that were incongruent with the movements of both partners (i.e., during same interactions) compared to their impact when they were congruent with the movement of the partner only (i.e., during opposite interactions); the index on GA captured the performance in opposite movements compared to same ones when visual cues appeared in spatially incongruent location relative to one's own movements. Thus, a positive correlation indicates that the larger the CCE, the greater the impairment in GA, suggesting that low-level intersensory effects and higher-order interpersonal motor abilities might

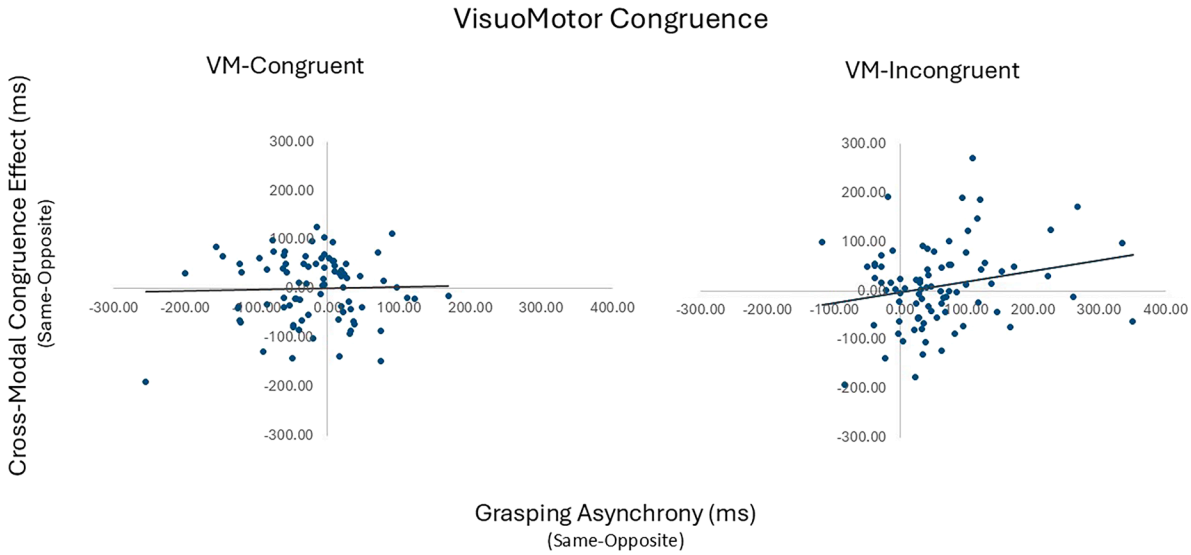


Figure 9. Positive correlation between cross-modal congruence effect during VisuoMotor incongruent (same—opposite movement type) in tactile detection task and grasping asynchrony in interactive conditions

Each dot represents a single participant's cross-modal interference effect. Solid black lines represent linear fits. Positive values correspond to increased cross-modal interference and grasping asynchrony.

be associated with one another. As a control analysis, the same index was computed on VisuoMotor congruent trials instead of incongruent ones (i.e., index = (interactive VisuoMotor congruent same movements)—(interactive VisuoMotor congruent opposite movements)) to evaluate whether the presentation of visual cues that were congruent with the movements of an individual but incongruent to those of the partner resulted in a facilitation in RTs and GA. Correlation significance was thus corrected for 2 comparisons (significant $p = 0.025$).

The correlation between participants' RTs to detect the tactile stimulation (i.e., the CCE in the tactile detection response) and their asynchrony in grasping the bottle together with the avatar (grasping asynchrony) was found to be significant only in interactive VM-incongruent conditions. Specifically, slower detection of visuo-tactile incongruent stimuli during same, compared to opposite, movements was associated with greater asynchrony in the joint grasping task ($r(84) = 0.22$; $p = 0.022$). Conversely, this result was not found for the VM-congruent interactive condition ($p = 0.81$) (Figure 9).

This indicated that participants who experienced less interference in the detection of tactile stimuli associated with visual cues that appeared in spatially incongruent locations during opposite movements also showed better synchronization with the partner in the same condition. This finding might suggest an impact of CCE on coordination performance in the interactive tasks, and that efficient tactile integration might be crucial for maintaining synchronized interactions.

DISCUSSION

Intersensory mechanisms are ubiquitous in humans' lives. They are fundamental to building a reliable representation of the environment and of one's own body schema, as well as for guiding our interaction with the surrounding world. While much of what is known about these mechanisms comes from studies concerning individuals dealing with non-social stimuli happening in the environment, the present studies aimed at describing the modulatory and relevant role of interpersonal interaction on these cross-modal mechanisms.

The present experiments were designed to investigate whether the integration of visual and tactile information during an interpersonal interaction is influenced by the nature of these interactions. Research has shown that visuo-tactile CCE, usually occurring for stimuli near the body,[12,32,46] are modulated by the fact that the individual prepares to and executes a movement so that CCE are observed even when the visual cue is distant from the body but aligned with the hand's endpoint position.[27–31,47,48] Importantly, this effect was also described during action observation.[31] However, if and how these intersensory effects change during interpersonal motor interactions, where individuals need to online predict their partner's movements, remains unexplored.

Across three experiments, we described that CCE were observed when a task-irrelevant visual cue appeared on the interacting partner's object and not one's own, only in conditions requiring continuous prediction and monitoring of the partners' movements (interactive condition). During joint motor interactions, it is known that the frontal performance monitoring system is attuned to others' actions,[49,50] enabling individuals to predict and evaluate their partners' ongoing behavior in real time. At a neural level, the tasks implemented in the present studies are supposed to involve fronto-parietal networks which are known to support (1) the planning and execution of upper-limb reaching and grasping movements,[51,52] (2) the ability to internally simulate observed actions of a partner,[53,54] (3) the ability to coordinate

with others in tasks requiring reaching and grasping movements,[40] and importantly, (4) the ability to integrate multisensory information in the peripersonal space.[55] Together, these mechanisms may provide a plausible neural substrate for the present findings. This suggests that when individuals engage in predictive coordination with others, the fronto-parietal circuits that normally integrate visuo-tactile information within the self's PPS may transiently incorporate sensory information from the partner's action space. This interpersonal modulation of CCE suggests that linking ones' own action to the partner's movements shapes the spatial rules governing PPS intersensory processing. This pattern is consistent with a dynamic change of the representations of one's own action-related sensory information and of the PPS, possibly supporting individuals' ability to facilitate sensorimotor predictions and consequently adapt one's own behavior to the movements of the partner. It should be stated that although the findings are interpreted in relation to fronto-parietal networks and peripersonal space mechanisms, the present experiments relied solely on behavioral measures, without neural data. Therefore, conclusions regarding the underlying neural mechanisms underpinning the found effects remain inferential.

All in all, our experimental results are in line with the previous literature on CCEs,[28–31] and suggest that multisensory integration is influenced by different characteristics of the interpersonal motor interactions during joint actions. Crucially, we extend this knowledge by highlighting that the ability to predict the movement of one's partner is critical for the emergence of these modulations. Conversely, no intersensory modulation was found when participants did not need to predict the movement of the partner (i.e., in the cued condition). Moreover, the task implemented here allowed us to investigate how sensory events are integrated based on the interaction itself rather than individual action goals.

These findings deepen our understanding of how multisensory information is integrated during joint action tasks, and future studies may focus on exploring how different interpersonal factors, personality traits, and fronto-parietal neural systems impact CCE during social exchanges, fostering our understanding of the interplay between low-level sensorimotor mechanisms and interpersonal processes. Overall, the results of the present study revealed an impact of interpersonal coordination on multisensory-motor processes and open new questions about the value of sociality in peripersonal space representation.

Limitations of the study

Some limitations should be considered when interpreting the present findings. First, the experiments relied exclusively on behavioral measures, so conclusions about the underlying neural mechanisms remain indirect. Although fronto-parietal and PPS networks provide a plausible framework for interpretation, neural activity was not directly assessed.

Second, the interaction paradigm involved coordination with a virtual partner rather than a physically co-present human partner acting on a shared object. Although the present task does not involve a real partner acting on a shared physical object, our working hypothesis is that in the interactive condition it captures a key component of interpersonal interaction—namely, the need to continuously predict and coordinate with a partner's movements in order to achieve a goal that cannot be accomplished alone. Consistent with this view, previous studies using this joint action paradigm have shown that coordination and synchronization with a virtual partner are sufficient to elicit behavioral interference effects.[41,42] These effects have been linked to frontal monitoring processes[49,56–58] and to left parietal regions involved in reaching movements.[51] Although similar effects have also been reported in versions of the task involving two real participants,[38,40,44,58,59] future studies using more immersive and physically interactive settings will be important to further assess the generalizability of these findings to real-world social interactions.

Finally, the study focused on a specific form of interpersonal coordination involving synchronous reach-to-grasp actions and visual-tactile stimuli delivered along the movement trajectory. It remains unclear whether similar effects would emerge in other social contexts, with different movement types, or other multisensory environments.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Matteo Candidi (matteo.candidi@uniroma1.it).

Materials availability

This study did not generate new materials.

Data and code availability

- All data reported in this paper will be shared by the lead contact upon request.
- This paper does not report original code.
- Any additional information required to re-analyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

The authors are grateful to all of the participants involved in the study and to Gokhan Karci for the support in the pilot recording.

Funding: M.C. was funded by Sapienza University (grant nos. RG123188B4631694 and RM124191116DD5CE). G.C. was funded by Avvio alla Ricerca 2022, from Sapienza University grant (grant no. AR12218 169299E74). V.E. was supported by Bando Giovani Ricercatori, Ministero italiano della Salute (grant no. GR-2021-12372923).

AUTHOR CONTRIBUTIONS

Conceptualization, G.C., A.A., U.P., V.E., and M.C.; methodology, G.C., U.P., V.E. and M.C.; software, G.C., U.P.; formal analysis, G.C., V.E. and M.C.; investigation, G.C. and A.A.; writing—original draft, G.C. and M.C.; writing—review and editing, G.C., A.A., U.P., V.E. and M.C.; supervision, V.E. and M.C.; funding acquisition, G.C., V. E. and M.C.

DECLARATION OF INTERESTS

The authors declare no competing interests.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work, the authors used generative AI tools (ChatGPT by OpenAI) to assist in language editing, improving clarity and refining phrasing in the manuscript. After using this tool, the authors reviewed and edited the content as needed, guaranteeing the integrity and accuracy of

the reported material and taking full responsibility for the content of the publication.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Human participants
- METHOD DETAILS
  - Common set-up
  - Visual stimuli of the avatar reaching and grasping movements
  - Visuo-tactile stimulation and tactile localization/detection task
  - Pilot study
  - Experiment 1
  - Experiment 2
  - Experiment 3
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2026.116950.

Received: December 1, 2025
Revised: April 13, 2026
Accepted: July 10, 2026

REFERENCES

1. Graziano, M.S., and Gross, C.G. (1995). The representation of extrapersonal space: A possible role for bimodal, visual-tactile neurons. In The cognitive neurosciences, M.S. Gazzaniga, ed. (MIT Press), pp. 1021–1034.
2. Rizzolatti, G., Fadiga, L., Fogassi, L., and Gallese, V. (1997). The space around us. Science *277*, 190–191. https://doi.org/10.1126/science.277.5323.190.
3. 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. https://doi.org/10.1016/s0028-3932(00)00150-0.
4. Maravita, A., Spence, C., and Driver, J. (2003). Multisensory integration and the body schema: close to hand and within reach. Curr. Biol. *13*, R531–R539. https://doi.org/10.1016/s0960-9822(03)00449-4.
5. Cooke, D.F., and Graziano, M.S.A. (2003). Defensive movements evoked by air puff in monkeys. J. Neurophysiol. *90*, 3317–3329. https://doi.org/10.1152/jn.00513.2003.
6. Serino, A., Noel, J.P., Galli, G., Canzoneri, E., Marmaroli, P., Lissek, H., and Blanke, O. (2015). Body part-centered and full body-centered peripersonal space representations. Sci. Rep. *5*, 18603. https://doi.org/10.1038/srep18603.
7. Rizzolatti, G., Scandolara, C., Matelli, M., and Gentilucci, M. (1981). Afferent properties of periarcuate neurons in macaque monkeys. I. Somatosensory responses. Behav. Brain Res. *2*, 125–146. https://doi.org/10.1016/0166-4328(81)90052-8.
8. Rizzolatti, G., Scandolara, C., Matelli, M., and Gentilucci, M. (1981). Afferent properties of periarcuate neurons in macque monkeys. II. Visual responses. Behav. Brain Res. *2*, 147–163. https://doi.org/10.1016/0166-4328(81)90053-x.
9. Farnè, A., Demattè, M.L., and Làdavas, E. (2005). Neuropsychological evidence of modular organization of the near peripersonal space. Neurology *65*, 1754–1758. https://doi.org/10.1212/01.wnl.0000187121.30480.09.
10. Brozzoli, C., Gentile, G., and Ehrsson, H.H. (2012). That's near my hand! Parietal and premotor coding of hand-centered space contributes to localization and self-attribution of the hand. J. Neurosci. *32*, 14573–14582. https://doi.org/10.1523/JNEUROSCI.2660-12.2012.
11. Brozzoli, C., Makin, T.R., Cardinali, L., Holmes, N.P., and Farne, A. (2011). Peripersonal space: a multisensory interface for body-object interactions. In The neural bases of multisensory processes, M.M. Murray and M.T. Wallace, eds. (Frontiers in Neuroscience).
12. Spence, C., Pavani, F., and Driver, J. (2004). Spatial constraints on visual-tactile cross-modal distractor congruency effects. Cogn. Affect. Behav. Neurosci. *4*, 148–169. https://doi.org/10.3758/cabn.4.2.148.
13. Spence, C., Pavani, F., and Driver, J. (1998). What crossing the hands can reveal about crossmodal links in spatial attention. Abstracts of the Psychonomic Society *3*, 13.
14. Pavani, F., Spence, C., and Driver, J. (2000). Visual capture of touch: Out-of-the-body experiences with rubber gloves. Psychol. Sci. *11*, 353–359. https://doi.org/10.1111/1467-9280.00270.
15. Spence, C., Ranson, J., and Driver, J. (2000). Cross-modal selective attention: On the difficulty of ignoring sounds at the locus of visual attention. Percept. Psychophys. *62*, 410–424. https://doi.org/10.3758/bf03205560.
16. Canzoneri, E., Magosso, E., and Serino, A. (2012). Dynamic sounds capture the boundaries of peripersonal space representation in humans. PLoS One *7*, e44306. https://doi.org/10.1371/journal.pone.0044306.
17. Graziano, M.S., Hu, X.T., and Gross, C.G. (1997). VisuoMotor properties of ventral premotor cortex. J. Neurophysiol. *77*, 2268–2292. https://doi.org/10.1152/jn.1997.77.5.2268.
18. Cléry, J., Guipponi, O., Odouard, S., Wardak, C., and Ben Hamed, S. (2015). Impact prediction by looming visual stimuli enhances tactile detection. J. Neurosci. *35*, 4179–4189. https://doi.org/10.1523/JNEUROSCI.3031-14.2015.
19. Cléry, J., Guipponi, O., Odouard, S., Pinède, S., Wardak, C., and Ben Hamed, S. (2017). The prediction of impact of a looming stimulus onto the body is subserved by multisensory integration mechanisms. J. Neurosci. *37*, 10656–10670. https://doi.org/10.1523/JNEUROSCI.0610-17.2017.
20. Fogassi, L., Gallese, V., di Pellegrino, G., Fadiga, L., Gentilucci, M., Luppino, G., Matelli, M., Pedotti, A., and Rizzolatti, G. (1992). Space coding by premotor cortex. Exp. Brain Res. *89*, 686–690. https://doi.org/10.1007/BF00229894.
21. 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. https://doi.org/10.1152/jn.1996.76.1.141.
22. Fogassi, L., Raos, V., Franchi, G., Gallese, V., Luppino, G., and Matelli, M. (1999). Visual responses in the dorsal premotor area F2 of the macaque monkey. Exp. Brain Res. *128*, 194–199. https://doi.org/10.1007/s002210050835.
23. Graziano, M.S., and Gross, C.G. (1993). A bimodal map of space: somatosensory receptive fields in the macaque putamen with corresponding visual receptive fields. Exp. Brain Res. *97*, 96–109. https://doi.org/10.1007/BF00228820.
24. Graziano, M.S., Yap, G.S., and Gross, C.G. (1994). Coding of visual space by premotor neurons. Science *266*, 1054–1057. https://doi.org/10.1126/science.7973661.
25. Ishida, H., Nakajima, K., Inase, M., and Murata, A. (2010). Shared mapping of own and others' bodies in visuotactile bimodal area of monkey parietal cortex. J. Cogn. Neurosci. *22*, 83–96. https://doi.org/10.1162/jocn.2009.21185.
26. 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. https://doi.org/10.1016/s0010-9452(08)70468-4.

27. Noel, J.P., Grivaz, P., Marmaroli, P., Lissek, H., Blanke, O., and Serino, A. (2015). Full body action remapping of peripersonal space: the case of walking. Neuropsychologia *70*, 375–384. https://doi.org/10.1016/j.neuropsychologia.2014.08.030.
28. Brozzoli, C., Pavani, F., Urquizar, C., Cardinali, L., and Farnè, A. (2009). Grasping actions remap peripersonal space. Neuroreport *20*, 913–917. https://doi.org/10.1097/WNR.0b013e32832c0b9b.
29. Senna, I., Cardinali, L., Farnè, A., and Brozzoli, C. (2019). Aim and plausibility of action chains remap peripersonal space. Front. Psychol. *10*, 1681. https://doi.org/10.3389/fpsyg.2019.01681.
30. Patané, I., Cardinali, L., Salemme, R., Pavani, F., Farnè, A., and Brozzoli, C. (2019). Action Planning Modulates Peripersonal Space. J. Cogn. Neurosci. *31*, 1141–1154. https://doi.org/10.1162/jocn_a_01349.
31. Patané, I., Brozzoli, C., Koun, E., Frassinetti, F., and Farnè, A. (2021). Me, you, and our object: Peripersonal space recruitment during executed and observed actions depends on object ownership. J. Exp. Psychol. Gen. *150*, 1410–1422. https://doi.org/10.1037/xge0001001.
32. Serino, A. (2019). Peripersonal space (PPS) as a multisensory interface between the individual and the environment, defining the space of the self. Neurosci. Biobehav. Rev. *99*, 138–159. https://doi.org/10.1016/j.neubiorev.2019.01.016.
33. Fossataro, C., Galigani, M., Sebastiano, A.R., Bruno, V., Ronga, I., and Garbarini, F. (2023). Spatial proximity to others induces plastic changes in the neural representation of the peripersonal space. iScience *26*, 105879.
34. Heed, T., Habets, B., Sebanz, N., and Knoblich, G. (2010). Others' Actions Reduce Crossmodal Integration in Peripersonal Space. Curr. Biol. *20*, 1345–1349. https://doi.org/10.1016/j.cub.2010.05.068.
35. Thomas, R., Press, C., and Haggard, P. (2006). Shared representations in body perception. Acta Psychol. *121*, 317–330. https://doi.org/10.1016/j.actpsy.2005.08.002.
36. Pezzulo, G., Iodice, P., Ferraina, S., and Kessler, K. (2013). Shared action spaces: a basis function framework for social re-calibration of sensorimotor representations supporting joint action. Front. Hum. Neurosci. *7*, 800. https://doi.org/10.3389/fnhum.2013.00800.
37. Fanghella, M., Era, V., and Candidi, M. (2021). Interpersonal motor interactions shape multisensory representations of the peripersonal space. Brain Sci. *11*, 255. https://doi.org/10.3390/brainsci11020255.
38. Candidi, M., Curioni, A., Donnarumma, F., Sacheli, L.M., and Pezzulo, G. (2015). Interactional leader–follower sensorimotor communication strategies during repetitive joint actions. J. R. Soc. Interface *12*, 20150644. https://doi.org/10.1098/rsif.2015.0644.
39. Candidi, M., Sacheli, L.M., Era, V., Canzano, L., Tieri, G., and Aglioti, S.M. (2017). Come together: human–avatar on-line interactions boost joint-action performance in apraxic patients. Soc. Cogn. Affect. Neurosci. *12*, 1793–1802. https://doi.org/10.1093/scan/nsx114.
40. Era, V., Candidi, M., Gandolfo, M., Sacheli, L.M., and Aglioti, S.M. (2018). Inhibition of left anterior intraparietal sulcus shows that mutual adjustment marks dyadic joint-actions in humans. Soc. Cogn. Affect. Neurosci. *13*, 492–500. https://doi.org/10.1093/scan/nsy022.
41. Era, V., Aglioti, S.M., Mancusi, C., and Candidi, M. (2020). Visuo-motor interference with a virtual partner is equally present in cooperative and competitive interactions. Psychol. Res. *84*, 810–822. https://doi.org/10.1007/s00426-018-1090-8.
42. Era, V., Candidi, M., Pezzetta, R., Pulcini, C., D'Antonio, S., Zabberoni, S., Peppe, A., Costa, A., Taglieri, S., Carlesimo, G.A., et al. (2023). The dopaminergic system supports flexible and rewarding dyadic motor interactive behaviour in Parkinson's Disease. Soc. Cogn. Affect. Neurosci. *18*, nsac040. https://doi.org/10.1093/scan/nsac040.
43. Fini, C., Era, V., Da Rold, F., Candidi, M., and Borghi, A.M. (2021). Abstract concepts in interaction: The need of others when guessing abstract concepts smooths dyadic motor interactions. R. Soc. Open Sci. *8*, 201205. https://doi.org/10.1098/rsos.201205.
44. Sacheli, L.M., Candidi, M., Pavone, E.F., Tidoni, E., and Aglioti, S.M. (2012). And yet they act together: Interpersonal perception modulates VisuoMotor interference and mutual adjustments during a joint-grasping task. PLoS One *7*, e50223. https://doi.org/10.1371/journal.pone.0050223.
45. Sacheli, L.M., Aglioti, S.M., and Candidi, M. (2015). Social cues to joint actions: The role of shared goals. Front. Psychol. *6*, 1034. https://doi.org/10.3389/fpsyg.2015.01034.
46. Spence, C., Nicholls, M.E., Gillespie, N., and Driver, J. (1998). Cross-modal links in exogenous covert spatial orienting between touch, audition, and vision. Percept. Psychophys. *60*, 544–557. https://doi.org/10.3758/bf03206045.
47. Brozzoli, C., Cardinali, L., Pavani, F., and Farnè, A. (2010). Action-specific remapping of peripersonal space. Neuropsychologia *49*, 3329–3332.
48. Lohmann, J., Belardinelli, A., and Butz, M.V. (2019). Hands ahead in mind and motion: Active inference in peripersonal hand space. Vision *3*, 15. https://doi.org/10.3390/vision3020015.
49. Moreau, Q., Tieri, G., Era, V., Aglioti, S.M., and Candidi, M. (2022). The performance monitoring system is attuned to others' actions during dyadic motor interactions. Cereb. Cortex *33*, 222–234. https://doi.org/10.1093/cercor/bhac063.
50. Era, V., Pesci, U.G., Moreau, Q., Pezzetta, R., Zabberoni, S., Peppe, A., Costa, A., Taglieri, S., Candidi, M., Aglioti, S.M., and Aglioti, S.M. (2025). Dopaminergic modulation of behavioral and electrocortical markers of interpersonal performance monitoring in Parkinson's Disease. Commun. Biol. *8*, 1195. https://doi.org/10.1038/s42003-025-08638-z.
51. Castiello, U. (2005). The neuroscience of grasping. Nat. Rev. Neurosci. *6*, 726–736. https://doi.org/10.1038/nrn1744.
52. Turrini, S., Bevacqua, N., Cataneo, A., Chiappini, E., Fiori, F., Candidi, M., and Avenanti, A. (2023). Transcranial cortico-cortical paired associative stimulation (ccPAS) over ventral premotor-motor pathways enhances action performance and corticomotor excitability in young adults more than in elderly adults. Front. Aging Neurosci. *15*, 1119508. https://doi.org/10.3389/fnagi.2023.1119508.
53. Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. Annu. Rev. Neurosci. *27*, 169–192. https://doi.org/10.1146/annurev.neuro.27.070203.144230.
54. Turrini, S., Fiori, F., Bevacqua, N., Saracini, C., Lucero, B., Candidi, M., and Avenanti, A. (2024). Spike-timing-dependent plasticity induction reveals dissociable supplementary–and premotor–motor pathways to automatic imitation. Proc. Natl. Acad. Sci. USA *121*, e2404925121. https://doi.org/10.1073/pnas.2404925121.
55. Serino, A., Canzoneri, E., and Avenanti, A. (2011). Fronto-parietal areas necessary for a multisensory representation of peripersonal space in humans: an rTMS study. J. Cogn. Neurosci. *23*, 2956–2967. https://doi.org/10.1162/jocn_a_00006.
56. Moreau, Q., Candidi, M., Era, V., Tieri, G., and Aglioti, S.M. (2020). Midline frontal and occipito-temporal activity during error monitoring in dyadic motor interactions. Cortex *127*, 131–149. https://doi.org/10.1016/j.cortex.2020.01.020.
57. Boukarras, S., Ozkan, D.G., Era, V., Moreau, Q., Tieri, G., and Candidi, M. (2022). Midfrontal theta transcranial alternating current stimulation facilitates motor coordination in dyadic human-avatar interactions. J. Cogn. Neurosci., 897–915. https://doi.org/10.1162/jcon_a_01834.
58. Boukarras, S., Era, V., Aglioti, S.M., and Candidi, M. (2021). Competence-based social status and implicit preference modulate the ability to coordinate during a joint grasping task. Sci. Rep. *11*, 5321. https://doi.org/10.1038/s41598-021-84280-z.
59. Curioni, A., Minio-Paluello, I., Sacheli, L.M., Candidi, M., and Aglioti, S.M. (2017). Autistic traits affect interpersonal motor coordination by

modulating strategic use of role-based behavior. Mol. Autism *8*, 23. https://doi.org/10.1186/s13229-017-0141-0.

60. Cohen, J. (2013). Statistical Power Analysis for the Behavioral Sciences (routledge).

61. Campbell, J.I.D., and Thompson, V.A. (2012). MorePower 6.0 for ANOVA with relational confidence intervals and Bayesian analysis. Behav. Res. Methods *44*, 1255–1265. https://doi.org/10.3758/s13428-012-0186-0.

62. Briggs, G.G., and Nebes, R.D. (1975). Patterns of hand preference in a student population. Cortex *11*, 230–238. https://doi.org/10.1016/s0010-9452(75)80005-0.

63. Fusaro, M., Tieri, G., and Aglioti, S.M. (2019). Influence of cognitive stance and physical perspective on subjective and autonomic reactivity to observed pain and pleasure: An Immersive Virtual Reality Study. Conscious. Cogn. *67*, 86–97. https://doi.org/10.1016/j.concog.2018.11.010.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
MorePower 6.0.4	More Power	RRID: SCR_024210
E-prime 2.0	E-prime	RRID: SCR_009567
Statistica	TIBCO	RRID: SCR_014213
JASP 0.18.3.0	Jasp-stats4	RRID: SCR_015823
R version 3.5.1	R software	RRID: SCR_001905
3DS Max 2017	Autodesk	RRID: SCR_014251

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Human participants

Written informed consent was obtained from all participants, in conformity to the ethical standards outlined in the Declaration of Helsinki. Participants were initially kept unaware of the study's objective and were informed only after completing all experimental procedures. All participants were caucasians and had either normal or corrected-to-normal vision. Potential effects of sex and gender were not evaluated in the present study, we acknowledge this as a limitation that may restrict the generalizability of the results. However, comprehensive clinical and demographic information for all participants is provided in Table S19. In Experiment 1, 15 euros were refunded for participating in the study. In Experiment 2 and 3 participation was not refunded. The study was approved by the Fondazione Santa Lucia of Rome (RM, Italy - Prot. CE/PROG.658).

Experiment 1

Reported effect sizes for differences between congruent and incongruent visuo-tactile multisensory effects in social and action tasks, range between ηp2 = 0.12–0.35.[29,31] Accordingly, we conducted a power analysis[60] using MorePower 6.0.4[61] assuming a $\eta_p^2 = 0.21$, $\alpha = 0.05$, power = 0.8. This yielded to an estimated sample size of 32 participants. A total of 36 individuals (19 females; mean age 25.66 years, ranging from 36 to 18 years old; S.D. = 3.14) participated in the study, but 2 of them were identified as outliers in the Tactile Localization/Detection task, while 3 of them in the Interpersonal Motor Task and subsequently excluded from all analyses (see statistical analysis section in the supplementary), resulting in a final sample of 34 participants for the Tactile Localization/Detection task and 33 for the Interpersonal Motortask. A total of 32 participants were confirmed as right-handed based on the Standard Handedness Inventory,[62] while the others were left-handed.

Experiment 2

In Experiment 2 and 3, the choice of the sample size was selected as well through a statistical power analysis[60] performed with the software More Power.[61] The expected effect size is the partial eta square ($\eta_p^2 = 0.29$) value, from Experiment 1. The output indicated that in a 2x2x2×2 within factors design, a power of 0.8 and an eta squared of 0.29, requires a sample size of 22 participants. In Experiment 2, a total of 28 individuals participated in the study, but 4 of them were identified as outliers and were subsequently excluded from all analyses. A final sample of 24 participants took part in the experiment. All participants (19 females; mean age 25.93 years, ranging from 20 to 35 years old; ±S.D. = 3.82) were initially unaware of the study's objective and were informed only after completing all experimental procedures. Twenty-one of them were right-handed, based on the Standard Handedness Inventory,[62] while the others were left-handed.

Experiment 3

In Experiment 3, a total of 20 individuals participated in the study, but 2 of them were identified as outliers, therefore subsequently excluded from all analyses (see statistical analysis section). All participants (8 females; mean age 24.85 years, ranging from 22 to 30 years old; S.D. = 2.58) were confirmed as right-handed based on the Standard Handedness Inventory.[62]

METHOD DETAILS

This section comprehensively outlines the methods shared across the three distinct experiments, providing an overview of the procedures adopted. In all experiments participants were involved in an interpersonal motor task while receiving tactile stimuli over their (moving) right hand finger(s). They were asked to synchronise their reach and grasp movements with those of a virtual partner in different conditions while their main task was to localize and/or detect as soon as possible the tactile stimulus. All experiments share the methods and procedures described below, except for specific details described in each experimental paragraph.

Common set-up

Interpersonal motor (reaching and grasping) task

We adapted a well-established human-avatar interactive task, that combined an ecological set-up with experimental control.[38–45] This task had previously demonstrated its ability to engage the same behavioral mechanisms observed in human-human interactions, specifically mutual adjustment and automatic imitation.[38,40,45,58,59]

Participants were seated in front of a rectangular table and an LCD monitor was positioned at the opposite edge of the table, 60 cm away from participants' eyes where the interaction partner was shown (see below). The task assigned to participants required them to reach and grasp a bottle-shaped object (37 cm in height) positioned next to the middle of the table, between them and the monitor and to touch the object in synchrony with their partner.

Participants were required to perform a reaching-to-grasp movement in synchrony with an avatar's movement. Considering the diameter of the bottle-shaped object, made of two rectangles in a stack with distinct widths, they could either perform a precision grip (on the upper part, 2.7 cm diameter) or a power grip (on the lower part, 6.5 cm diameter) with their right hand. Prior to each trial, participants were instructed to press and hold, with their index finger and thumb, a start button located on the right side of their midline (at 34 cm distance from the bottle object and 10 cm away from the midline). Following an auditory signal played through headphones, they released the button to initiate the grasping movement together with the virtual avatar presented on the screen in front of them. Grasping times were measured using touch-sensitive markers placed in pairs on each rectangle, precisely at distances of 15 cm and 22 cm along the vertical height of the object.

Two different auditory stimuli were used to instruct participants in the task, depending on the experimental mode of interaction. Indeed, the experimental task involved two conditions that differ in the interactivity level (Interaction Type Factor): (1) in the Interactive condition, the goal of participants' action cannot be achieved without considering that of the virtual partner, taking into account its movements and adapting to them in real time; (2) in the Cued condition, participants do not need to coordinate their behavior to that of the partner, while grasping in synchrony with it. One auditory stimulus was associated to the Interactive condition in which participants were instructed to perform opposite (e.g., if the virtual partner grasped the lower bottle part, the participant grasped the upper part, and vice versa) or same (e.g., both grasping the upper or lower part of their bottles) movements in relation to the avatar. Before the commencement of each trial, participants received audio instructions in the form of vocal commands, specifying whether they had to perform same or opposite movements. Conversely, the auditory stimulus associated with the Cued condition, instructed participants that they would hear one of two tones indicating whether they should grasp the upper or lower part of the bottle. A high-pitched tone (1470 Hz) signaled to reach toward the upper white panels, while a lower-pitched tone (114 Hz) indicated reaching toward the lower white panels. Each of the four auditory cues had a duration of approximately 190 ms. Prior to initiating different conditions, participants were instructed on the corresponding auditory cues. To ensure participants' comprehension of the instructions, a set of try-outs was conducted before each condition. Once the trial runs were successfully completed, the actual experiment began and the starting condition was randomised and counterbalanced between participants. In both Cued/Interactive conditions, the final participants' goal was to grasp the bottle-shaped object as synchronously as possible with the virtual partner. For both interaction modes, the movements, with respect to the partner's grip, could either be Same or Opposite (Interaction Type Factor); depending on whether they grasped the bottle with the same grip or in a complementary manner relative to the avatar. The details of the setup are mainly illustrated in the main text (Figure 1A).

The trial timeline was as follows: participants were presented with a fixation cross lasting for 300 ms, followed by an auditory instruction via headphones, either the Interactive (Same/Opposite) or Cued (Up/Down). Afterward, the Avatar's appearance on the screen (see visual stimuli of the avatar reaching and grasping movements section) marked the soon beginning of its movement, which occurred between 200 ms and 600 ms. Upon receiving the auditory instruction and seeing the avatar, participants were allowed to release the start button and synchronously move to grasp the bottle-shaped object along with the avatar. The Avatar executed its action during a trial duration of approximately 2000 ms. The graphical example of a trial timeline can be found in the main text (Figure 1B).

The time between trials was determined by participants' return to the starting position, with the subsequent trial initiated by the experimenter upon their return on the start button. Trials where participants initiated their movement prior to the auditory instruction were excluded from the analysis. A valid trial denoted participants' adherence to the auditory instructions, accurate execution of the bottle grasp and detection of the tactile stimulation presented. Presentation and randomization of the stimuli were administered through E-Prime2 Professional software (Psychology Software Tools Inc.).

Visual stimuli of the avatar reaching and grasping movements

As in previous studies[39,41,42,45,56] the kinematic attributes of the virtual partner were derived from the movements of human participants engaged in different grasping actions within a joint-grasping task involving human-human interactions (refer to Fusaro et al.[63] for comprehensive technical information about Motion Capture recording). Utilising MotionBuilder 2017 and 3DS Max 2017 (Autodesk) software, the processed trajectories were accurately generated and applied to a male character of Caucasian ethnicity. Hence, the stimuli represented a male avatar conducting grasping movements toward a bottle-shaped object, as the one presented to the participant in the experimental set-up.

The avatar featured solely the upper body, spanning from the shoulders downward, intentionally excluding the neck and head to avoid diverting subjects' attention to the virtual partner's facial expressions. A total of 10 distinct grasping motions were captured in

the video clips dataset. Half of these motions ended with the avatar's hand gripping the lower section of the bottle-shaped object (power grip), while the remaining half concluded with the avatar's hand grasping the upper portion (precision grip). An additional 4 clips were crafted, depicting digital avatars executing movement adjustments by transitioning from precision to power grip (or vice versa) during the reaching phase. These modifications were applied in 20% of the trials wherein online corrections were involved. The process of crafting correction videos involved combining the initial frames of one clip, like a power grasp clip, with the final frames of another, such as a precision grasp clip, accomplished using 3DS Max. These modified clips were considered as catch trials, so not relevant for the analysis, but to keep the participants engaged in the task. All video stimuli had a duration of approximately 2000 ms.

The grasp times of the avatar were logged on a trial-by-trial basis using a photodiode situated on the screen. These two photodiodes placed on the screen transmitted their signals to the tactile stimulator and to a computer that was then recorded their timing by E-Prime 2 Professional (version 2.0.10.242, Psychology Software Tools Inc., Pittsburgh, PA) through a TriggerStation (BrainTrends Ltd., Italy). The activation of one photodiode was initiated by a white square positioned at the bottom left of the screen (unseen by the participants) to trigger the tactile stimulation. Additionally, a second photodiode was activated by another white square placed at the bottom right of the screen (not visible to the participants), which coincided with the moment when the avatar virtually grasped its object.

Visuo-tactile stimulation and tactile localization/detection task

While performing the grasping task, participants were instructed to detect (and in Experiment 1 discriminate the location of) a tactile stimulation. Tactile stimulation was administered to participants using a constant-current electrical stimulator (DSA7A, Digitimer Ltd., Welwyn Garden City, United Kingdom). Disposable electrodes were positioned dorsally on both the index and the thumb (Experiment 1) or the solely index finger (Experiment 2, 3) of the participants' right hand, with which they performed the interpersonal motor task (see Interpersonal Motor (Reaching and Grasping) Task section), maintaining a 1 cm separation between them. To ascertain stimulus intensity, a tactile threshold procedure (Serino et al.[55]) was employed. Starting with minimal levels of tactile stimulation (100–400 VMax), the intensity was incrementally increased until participants reported a clear sensation. Following the initial reported level of tactile perception, participants were stimulated 10 times and requested to report whether they felt the stimulation. If participants successfully detected it 90% of the time (9 out of 10 instances), that value was adopted as the tactile stimulation threshold for the rest of the experiment. The duration of the tactile stimulus was 200μs. During the experiments, participants were asked to detect the tactile stimulus by executing a keyboard (Experiment 1) or foot-pedal response (Experiment 2).

Synchronously to the tactile stimulation, a visual cue appeared for ≈80 ms on the bottle of the partner in a position Congruent or Incongruent with respect to the final prehension movement of the participants on their own object (VisuoMotor Congruence factor; Figure 2A). In Experiment 1, since both index and thumb fingers were tactilely stimulated and the visual cue could appear in a Congruent or Incongruent position relative to the participant final finger position on its bottle (front or back), we measured another interaction effect with respect to the location of the visual cue and the participant end-finger position (Somatotopic Congruence factor; Figure 2B). In Experiment 2, the visual cue was presented with or without the tactile stimulus (Sensory Condition factor). In Experiment 3, the visuo-tactile stimulation was presented early or late (see below), relative to the avatar's movement start (Visuo-Tactile Timing).

The synchronization of the visuo-tactile stimulation was controlled and administered using E-Prime 2.0. (version 2.0.10.242, Psychology Software Tools Inc., Pittsburgh, PA). The precise timing of the tactile stimulus relative to the visual cue was verified using a photodiode positioned on the monitor, ensuring that the tactile stimulus was delivered exactly when the visual cue appeared on the screen (as described in the visual stimuli of the avatar reaching and grasping movements section). Additionally, the temporal alignment of the two stimuli was validated through a dedicated trial run. In this trial, beyond the photodiode verification, two cutaneous electrodes assessed whether the skin response to the tactile stimulus corresponded precisely to the timing of the tactile trigger using the gTec g.HIAMP Amplifier (g.Tec medical Engineering GmbH, Austria).

Pilot study

Prior to conducting the first experiment, a pilot was run in order to determine the timings within the video clips (see stimuli section) when participants could accurately predict the aimed location of the virtual partner's movement. In this context, video clips were presented to a group of 11 healthy individuals (6 Females, mean age 25.73 ± 2.11 SD years), with the clips halted at various stages within the reach-to-grasp action (1/5, 2/5, 3/5, and 5/5 of the total movement duration). Participants were prompted to respond by pressing one of three buttons (uncertain/up/down) to indicate whether, in their opinion, the virtual avatar aimed to grasp the bottle-shaped-object on the upper or lower part or that they were uncertain about the movement's target location. The findings showed that when the stimulus appeared in the first fifth of the reaching-to-grasp movement participants were unable to determine the avatar's target direction. In the second fifth of the movement duration, participants started to understand the movement direction, but the error rate was still high. Hence, drawing insight from the data (Figure S1), to be sure that participants could accurately discriminate the aimed location of the observed action, visual and tactile stimulation (see stimuli and procedure sections) during the actual experiment has been synchronised at the third-fifth of the overall movement duration.

Experiment 1

In this version of the human-avatar joint action task, experimental subjects were required to detect and discriminate as quickly as possible where the tactile stimulation occurred (right hand thumb/index), while ignoring visual distractors. To report their answer,

participants were instructed to use a keyboard, specifically: their left index finger to press the left arrow key (to indicate a tactile stimulus on their right index finger) or their left thumb to press the right arrow key (to indicate a tactile stimulus on their right thumb finger), depending on the location where they felt the stimulation on the other hand. The visual cue was presented at the corresponding final position of the partner's right index finger (on the front of the avatar's bottle) or thumb (on the back of the avatar's object). The visual cue might appear slightly to the left of the virtual object, corresponding to the avatar's index finger on either the upper or lower portion of the bottle. Alternatively, a yellow half-dot could appear slightly to the right of the bottle, aligning with the avatar's thumb on either the upper or lower portion of the bottle. The distinction between a full dot and a half dot was intended to convey the impression that the dots were either appearing on the front (full dot) or on the back of the virtual bottle (half dot).

The entire experiment comprised 480 trials, equally distributed between Interactive (240 trials) and Cued (240 trials) conditions, further divided between Opposite (120 trials) and Same (120 trials) movements. Within these, a subdivision was made between VisuoMotor Incongruent (60 trials) and VisuoMotor Congruent (60 trials) movements, reflecting the congruence between the participant's grasping target location and the location of the yellow dot/half-dot along the bottle's vertical axis (upper or lower portion). Moreover, there was an additional subdivision of these trials into Somatotopic Congruent (30 trials) and Incongruent (30 trials) stimulations, indicating the congruence between finger stimulation (index/thumb) and the dot's position along the bottle's depth axis (full/half dot). The 30 trials included Precision (10 trials) and Power (10 trials) grasping movements, and 10 different catch-trials: 6 trials with visuo-tactile cues delivered at random time points within six equidistant time intervals between 1/5th and 3/5th of the avatar's movement; 2 trials involving avatar course correction; 2 trials and with visual cues presented in the absence of tactile stimulation.

Experiment 2

The experiment consisted in a modified version of the task employed in Experiment 1, with two main condition: (1) a Visuo-Tactile bimodal condition, in which the tactile stimulation, delivered on the dorsal part of participants' right index finger, was synchronised with a visual cue (a full yellow dot) appearing on the upper or lower part of avatar's bottle-shaped object; (2) a Visual unimodal condition, in which the visual cue would be presented in absence of any tactile stimulation (no white square would be presented on the bottom left of the screen to trigger the stimulation; see stimuli section). To respond to either the visual cue or the tactile stimulation, participants were required to press a foot pedal positioned and secured beneath the table with their right foot. Participants were required only to perform a detection task, since the thumb stimulation was not present and only the full yellow dot appeared on the front of the avatar's bottle (in the upper/lower part) relative to Experiment 1.

The experiment comprised a total of 192 trials per Condition (Visuo-Tactile and Visual). These trials were equally divided between Interactive (96 trials) and Cued (96 trials) conditions. Within these trials, half of them (48 each) required to perform the Same movement as the avatar, while the remaining half (48 each) the Opposite. Furthermore, half of these 48 trials showed a visual cue that was Congruent (24) and Incongruent (24) with the location of the movement target. Among the 24 VisuoMotor Congruent and Incongruent trials, 8 exhibited a power grip, 8 a precision grip, 2 presented modified trajectories with avatar direction changes, and 6 catch trials in which the visuo-tactile stimulation was delivered at six random intervals between the 1/5 and the 3/5 percentage of the avatar's movement start.

Experiment 3

The procedure was identical to Experiment 1 and 2, except for the following changes. During the whole experiment participants would receive both the tactile and visual stimulation, except for the 10% of catch trials in which only the visual cue was presented and were not considered for the analyses. The tactile stimuli and response were administered as in the visuo-tactile sensory condition of Experiment 2. Based on the pilot results (see pilot section), the trials were equally differentiated in two visuo-tactile stimulation timings: 1) Early presentation, before the target location of the avatar's movement could be discriminated, corresponding to 1/5 of the avatar's movement; and 2) Late presentation, when the avatars' movement direction could be accurately predicted, corresponding to 3/5 of the avatar's movement.

The experiment comprised a total of 336 trials. These trials were equally divided between Interactive (168 trials) and Cued (168 trials) conditions. Within these trials, half of them (84 each) required to perform the Same movement as the avatar, while the remaining half (84 each) the Opposite. Furthermore, half were VisuoMotor Congruent (42) and half were Incongruent (42) trials within the Same and Opposite movement trials. Among these 42 trials, half presented an Early stimulation (21) and the other half a Late stimuli presentation. Within these 21 trials, 8 exhibited a power grip, 8 a precision grip, 2 presented modified trajectories with avatar direction changes, and 3 trials were catch trials without any tactile stimulation.

QUANTIFICATION AND STATISTICAL ANALYSIS

The dependent variables employed for the statistical analysis in the Tactile Localization/Detection Task comprises the Reaction Times (RTs) for the tactile response (in ms), corresponding to the timing between the appearance of the visuo-tactile stimulus and the localization (Experiment 1) or detection (Experiment 2 and 3) response of the participant to the tactile stimulus, measuring CCE in the paradigm.

For the Interpersonal Motor Task, we considered as dependent variables the Grasping Asynchrony (in ms), corresponding to the absolute value of the difference between the grasping time of the avatar and the participant on their own object, measuring the asynchrony of participants' grasp compared to the partner and the overall ability to perform the task.

For both tasks a within-subject Analysis of Variance (ANOVA) was performed, since assumptions for parametric testing, tested through the Shapiro-Wilks and Lilliefors tests, were generally met.

Error Rate was considered as the incorrect or missed identification of the tactile stimulus (localization in Experiment 1; detection in Experiments 2 and 3), limited to trials with correct joint grasp execution (range 0–1), hence, we employed nonparametric tests, specifically an Aligned Rank Transform (ART) procedure, where applicable or a Friedman ANOVA, to assess it.

For all our dependent variables, trials where participants missed the tactile stimulation (i.e., did not initiate keyboard/pedal response) and/or performed anticipatory responses. Furthermore, we excluded trials in which participants failed to grasp the touch-sensitive copper plates (resulting in no recorded response), released the start button before the go instruction, or the opposite/same instructions were not followed. Only data from accurate trials were taken into consideration for the analyses. At the individual trials level, any values exceeding or falling below ±2.5 standard deviations of the individual mean for each experimental condition were treated as outliers and excluded (on average, excluded trial 16% ± 2.16% of the total). At the group level, participants exhibiting values that diverged by ± 2.5 standard deviations above or below from the group mean in each experimental condition were identified as outliers and omitted from the analyses (on average, excluded trial 10% ± 3.26% of the total). All significance tests were conducted with an α significance level of 0.05, unless otherwise specified. Post-hoc tests, when applicable, were carried out using the Bonferroni test method.

To examine the relationship between performance in the interpersonal motor and tactile detection task, we conducted correlations between the RTs of the CCE and Grasping Asynchrony for the grasping task. To increase the statistical power of this correlation, we pooled data from the three studies, including only the data from the common factors of Interaction Mode (Cued/Interactive), Movement Type (Same/Opposite) and VisuoMotor Congruence (VM-Congruent/VM-Incongruent). Specifically: 1) In the first study, only the trials in which the visual cue was presented on the front portion of the bottle (full yellow dot); 2) In the second study, only the data from the multisensory Visuo-Tactile condition; 3) In the third study, only the trials from the Late Timing condition were considered. We then created an index by subtracting the mean values of all Opposite interaction trials from Same ones in Interactive VM-Incongruent and VM-Congruent conditions. These indices allowed us to assess whether the modulation of visuo-tactile interference effects elicited when coordinating with a partner to perform opposite compared to same movements was also associated with the ability to synchronously grasp the object during the interpersonal motor task.

Frequentist statistical analyses were conducted using Statistica software (TIBCO Software Inc) and R (version 3.5.1), while JASP 0.18.3.0 (JASP stats) was used for Bayesian Analyses and the graphs. Data, code, and materials are accessible upon request. In the results sections only the significant effects of interests are extensively reported, while the other results are showed in tables in the Supplementary. For all experiments, n refers to the number of participants included in the final analyses after exclusion of outliers. Specifically, $n = 34$ for the Tactile Localization Task and $n = 33$ for the Interpersonal Motor Task in Experiment 1, $n = 24$ in Experiment 2, and $n = 18$ in Experiment 3. Descriptive statistics are reported as mean ± standard deviation (SD), unless otherwise specified. Effect sizes are reported as partial eta squared (ηp^2) for ANOVAs and Bayes Factors (BF10) for Bayesian analyses. Exact statistical results, including test statistics, degrees of freedom, p values, and effect sizes, are reported in the Results section and in the corresponding Supplementary Tables. Figure legends specify the sample size, summary statistics displayed, and error bars where applicable. Asterisks indicate significant pairwise comparisons identified by Bonferroni-corrected post-hoc tests ($*p < 0.05$).