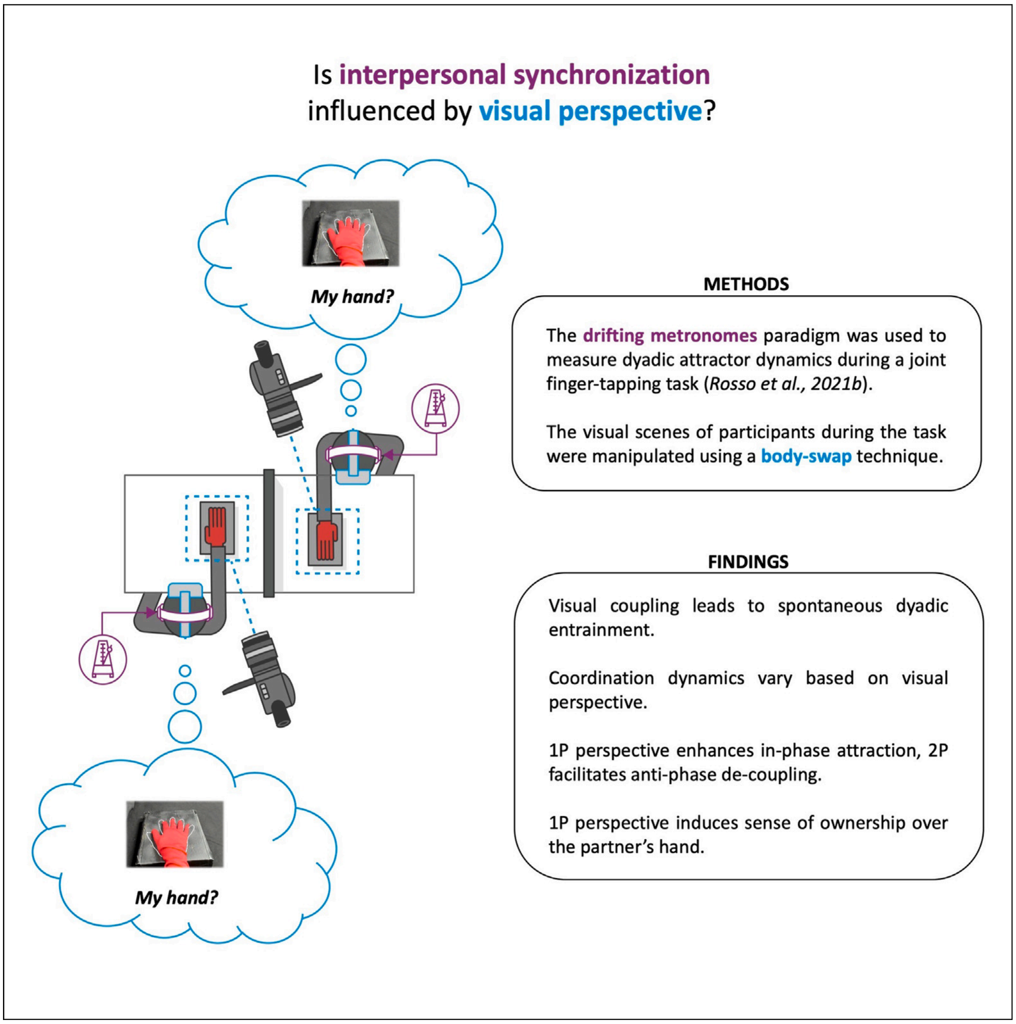
iScience

CellPress OPEN ACCESS

Article

Embodied perspective-taking enhances interpersonal synchronization: A body-swap study



Mattia Rosso, Bavo van Kerrebroeck, Pieter-Jan Maes, Marc Leman

mattia.rosso@ugent.be

Highlights

Visual coupling leads to spontaneous dyadic entrainment

Coordination dynamics vary based on visual perspective

1P perspective enhances attraction, while 2P perspective facilitates de-coupling

1P perspective induces sense of ownership over the partner's hand

Rosso et al., iScience *26*, 108099
November 17, 2023
https://doi.org/10.1016/j.isci.2023.108099

iScience

CellPress
OPEN ACCESS

Article

Embodied perspective-taking enhances interpersonal synchronization: A body-swap study

Mattia Rosso,[1,2,5,*] Bavo van Kerrebroeck,[1,3,4] Pieter-Jan Maes,[1] and Marc Leman[1]

SUMMARY

Humans exhibit a strong tendency to synchronize movements with each other, with visual perspective potentially influencing interpersonal synchronization. By manipulating the visual scenes of participants engaged in a joint finger-tapping task, we examined the effects of 1st person and 2nd person visual perspectives on their coordination dynamics. We hypothesized that perceiving the partner's movements from their 1st person perspective would enhance spontaneous interpersonal synchronization, potentially mediated by the embodiment of the partner's hand. We observed significant differences in attractor dynamics across visual perspectives. Specifically, participants in 1st person coupling were unable to maintain de-coupled trajectories as effectively as in 2nd person coupling. Our findings suggest that visual perspective influences coordination dynamics in dyadic interactions, engaging error-correction mechanisms in individual brains as they integrate the partner's hand into their body representation. Our results have the potential to inform the development of applications for motor training and rehabilitation.

INTRODUCTION

Humans exhibit a compelling tendency to synchronize rhythmic movements with one another.[1–4] As soon as two individuals exchange information via one or multiple sensory channels,[5] such phenomenon may occur spontaneously and even against the intention to ignore the other.[6] Visually mediated interactions, in particular, are governed by attractor dynamics[7,8] which stabilize dyadic behavior in recurrent and stable coordinative patterns, and are characterized by a dynamic balance between the pursuit of individual behavioral trajectories (*competition* process) and the attraction into coupled behavior (*cooperation* process).[6,9]

Ecological dyadic interactions take place in settings where individuals perceive each other from a face-to-face 2nd person perspective. However, embodied simulation accounts of social cognition posit that the mirroring of another person's movements is enabled by neural representations based on a bodily format,[10] which require a visuospatial transformation to remap the observed movement into an egocentric frame of reference.[11] Despite broad evidence for such form of embodied perspective-taking,[12–17] the mechanism remains overlooked in the literature on dyadic interactions.[11] In the present work, we investigate the role of visual perspective in temporal coordination, under the hypothesis that perceiving the movements of a partner from their 1st person perspective would facilitate spontaneous motor alignment, thereby enhancing interpersonal synchronization.

Perspective-taking can nowadays be induced in an embodied bottom-up fashion, by experimentally transposing the visual scenes perceived by two individuals into the partner's egocentric frame of reference.[18] Manipulations of this kind tap into the plasticity of body schemas as represented in the central nervous system,[19,20] and effectively lead to experience embodiment[21–24] and agency[25,26] over effectors not belonging to one's own body, as long as they are visually perceived in a configuration which is coherent with bodily constraints.[27] The same principle, originally investigated by means of the rubber hand illusion,[21] was extended to the more radical experience of full-body[18,23,28–30] and out-of-body[31,32] illusions, where a person gets to experience ownership and agency over a humanoid virtual avatar in 1st person, or dislocation respect to the position of the real body, respectively. Applied in a social setting, the same principles allow to induce a full body-swap between two real persons by streaming the 1st person view of one partner into the visual scene of the other.[33]

Koban et al.[34] proposed that dyadic synchronized behavior is guided by an optimization principle, aimed at minimizing prediction errors by correcting the temporal mismatch between movements executed by one self and movements executed by the other. In terms of brain-body-environment system, motor control is guided by environmental contingencies toward a reduction of computational cost.[35] With these principles in mind, let us take the human hand as paradigmatic effector to investigate embodiment,[21] and joint finger-tapping as paradigmatic task to investigate interpersonal synchronization.[6,36,37] Provided the hand of a partner can be integrated in one's own body schema when visually perceived in 1st person during joint finger-tapping,[38] we expected temporal mismatches to carry more weight as compared to the ecological

[1]IPEM - Institute for Systematic Musicology, Ghent University, 9000 Ghent, Flanders, Belgium
[2]PSITEC - Psychologie: Interactions, Temps, Emotions, Cognition - ULR 4072, University of Lille, 59650 Lille, Hauts-de-France, France
[3]SPL - Sequence Production Lab, McGill University, Montreal, Quebec H3A 1B1, Canada
[4]IDMIL – Input Devices. And Music Interaction Laboratory, McGill University, Montréal, Québec H3A 1E3, Canada
[5]Lead contact
*Correspondence: mattia.rosso@ugent.be
https://doi.org/10.1016/j.isci.2023.108099

2^{nd} person perspective, which would lead the partners to engage in a stronger error-correction response via dyadic entrainment. This is because, in such scenario, the other's hand is perceived as an embodied effector and therefore represented by the motor system in terms of motor potentiality for actions.[39,40] When the visual feedback expected from a motor output is altered, resulting in a mismatch with the prediction of a forward-model,[41–43] humans spontaneously engage in motor adaptation to keep a consistent relationship between action and perception.[44] This brings us to our central research questions. (1) Can we induce spontaneous interpersonal synchronization by visually coupling, in 1^{st} person perspective, two individuals engaged in a joint-finger tapping task? (2) Can visual coupling in 1^{st} person strengthen interpersonal synchronization, as compared to the ecological 2^{nd} person perspective? (3) How do attractor dynamics compare across different visual perspectives?

In order to answer these questions, we adopted the *drifting metronomes* paradigm for dyadic entrainment[6] and manipulated the visual scenes perceived by the partners across experimental conditions. The paradigm consists of a dyadic finger-tapping task where each partner is instructed to synchronize their taps with a different auditory metronome, while being visually exposed to the partner's hand. Crucially, the two metronomes are set at slightly different frequencies, such that participants are constantly exposed to the incongruent rhythm of the partner's finger-taps while attempting to maintain their instructed rhythm. With this expedient, the paradigm allows us to study spontaneous tendency to synchronize with one another (cooperation process) when individual intended rhythms are at odds (competition process). The cooperation process attracting individual timings away from the intended rhythm is the main object of our investigation. Previous work adopting this paradigm showed that, despite the explicit instruction to ignore each other, recurrent patterns of spontaneous coordinated behavior emerge between individuals according to consistent temporal dynamics.[6]

Another essential feature of the paradigm is a minimal gap in the frequencies of the two metronomes, such that when they are set to start at the same time, their relative phase systematically increases with every beat from 0 to π radians and subsequently decreases from π to 0. Cyclical repetitions of this pattern allow us to identify regions of maximal attraction over the whole attractor landscape,[45] capturing the time-varying nature of dyadic entrainment beyond a global measure of synchronization. The same task was performed under different conditions of body-swap. As illustrated in Figure 1, during the task participants were either seeing the other's hand in 2^{nd} person (1), their own hand in 1^{st} person (2), the other's hand in 1^{st} person (3) or their own hand in 2^{nd} person (4). The subjective feeling of embodiment was measured via a visuotactile stimulation procedure based on the principles of the rubber-hand illusion. Experimental design and procedures are described in detail in the STAR Methods section.

Finally, we tested the well-documented association between interpersonal synchronization and empathic traits (for a recent review, see[46]), and with the self-reported sense of ownership over the other's hand. At a higher cognitive level, a bottom-up driven experience of being "in the shoes of the other" mitigates outgroup[47] and racial biases,[48,49] attenuates gender stereotype threat,[50,51] promotes perceived self-other similarity[52,53] and even the social acceptability of a humanoid robot.[54] Crucially, both empathy and synchronization activate embodied representations of observed actions in the brain,[55] while high trait empathy[56] and empathic perspective-taking[57] were shown to strengthen such representations. We therefore hypothesized that both high scores in cognitive perspective-taking and subjective experience of embodiment would predict stronger entrainment with the partner, in particular when assuming their 1^{st} person visual perspective.

RESULTS

The results presented in the first part of this section focus on dyadic behavior. Specifically, we aimed to quantify the extent to which partners were temporally coordinated throughout the task. To achieve this, we based our examination on Joint Recurrence Quantification Analysis (JRQA),[58] replicating the implementation presented in Rosso et al.[6]. JRQA provided a recurrence score, which is a relational measure quantifying the degree of temporal coordination between the partners.[9] Notably, our JRQA implementation tracked the evolution of the recurrence score over the cycle of the drifting metronomes, illustrating the attractor landscape of each dyad.[45] In simpler terms, this allowed us to pinpoint attractor points throughout the cycle and evaluate their influence based on variations in the recurrence score.

For each pair of participants (dyad), we calculated recurrence scores for the drifting metronomes' cycles (ranging from 0 to 2π) and subsequently averaged these scores across 10 cycles for each experimental condition. Thus, the average recurrence score time series served as our response variable. By means of growth curve analysis,[59] we modeled the time series with 2^{nd} order orthogonal polynomials of Time. The full model included two two-level factors: Coupling (Coupled, Uncoupled) and Perspective (1P, 2P). With this approach, we were able to capture the temporal components of the recurrence score within the polynomial terms of the model (β_0 for the average, β_1 for the linear trend, and β_2 for the depth of the parabolic curvature), and analyze how these components were affected by our experimental manipulations of Coupling and Perspective. The Uncoupled conditions (control) were used as a baseline to test the effect of Coupling, whereas 2P conditions were used as baseline for testing the effect of Perspective.

In line with our hypotheses, we found a significant main effect of Coupling (*Estimate* = 436.278, *SE* = 73.777, $p < 0.001$) indicating an overall increase of the recurrence score in presence of informational coupling as compared to the uncoupled control conditions, independently from the manipulation of Perspective. We also found a significant interaction effect between Coupling and the quadratic term of Time (*Estimate* = 1201.669, *SE* = 166.386, $p < 0.001$), meaning that the modulation of the attractor landscape on the response variable resulted in a significant "valley" around the anti-phase midpoint in coupled conditions. Crucially, we found a 3-way interaction between the linear component of Time, Coupling and Perspective (*Estimate* = 478.850, *SE* = 215.621, $p = 0.026$), indicating that the linear coefficient in coupled conditions significantly differed across 1^{st} person and 2^{nd} person levels of Perspective. Such interaction captures the change in the asymmetry of the parabolic curves across the two Coupled conditions. Whilst in 2^{nd} person the recurrence score lingers in the "valley" into the second half of the drifting metronomes' cycle, in 1^{st} person it bottoms at the anti-phase point and grows straight toward the in-phase point. Figure 2 shows the grand-average curves of the recurrence score across experimental conditions. Table 1 shows the fixed effects parameter estimates

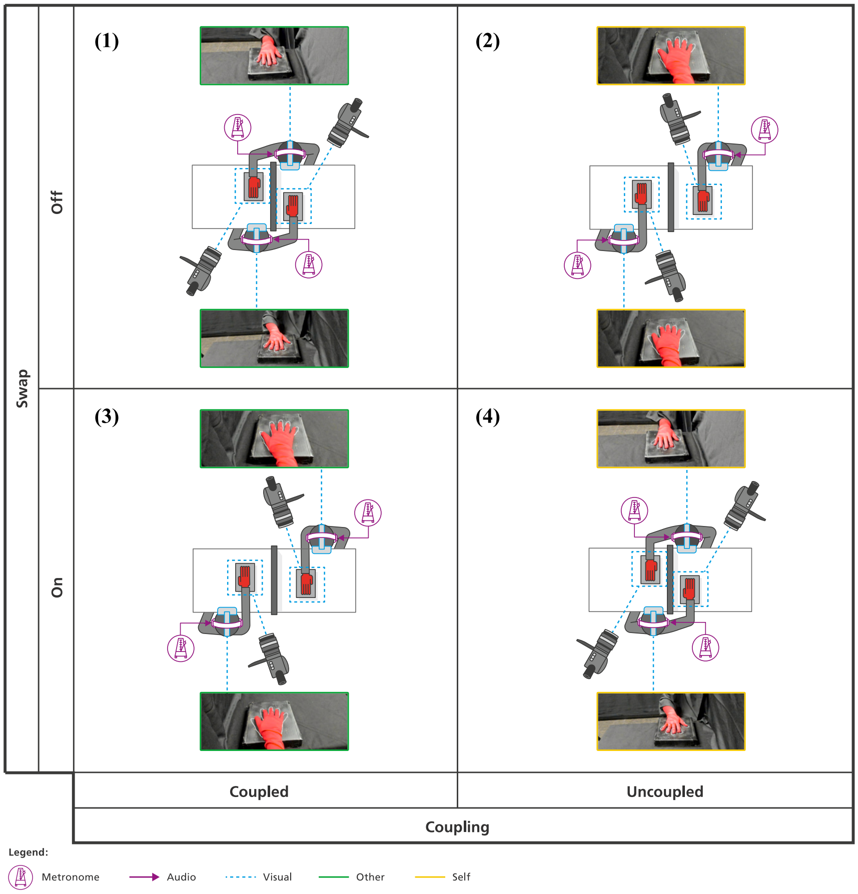


Figure 1. Experimental design

The study was designed in a Perspective (2P, 1P) x Coupling (Coupled, Uncoupled) factorial structure. Each participant was equipped with a headset providing full immersion in different visual scenes across experimental conditions. Visual scenes were captured and streamed in real-time by cameras placed either in front of the partner's hand or above the participant's shoulder, as illustrated in the detail boxes of the figure. The setup allowed for the crucial manipulation of *swapping* the visual scenes as captured by different angles, illustrated in the schema as "Swap On/Off." The design resulted in the following experimental conditions. ***Condition 1. "2P Coupled."*** Participants tapped along with an auditory metronome, while looking at the partner's hand tapping. The partner's hand was video-recorded from a frontal position and no swapping was performed, so that the hand was perceived from a 2nd person perspective. Participants were explicitly asked to neglect the partner's movements and focus on following their own metronome. ***Condition 2. "1P Uncoupled" (control).*** Participants tapped along with an auditory metronome, while looking at their own hand. Their own hand was video-recorded from above their shoulder and no swapping was performed, so that the hand was perceived from a 1st person perspective. ***Condition 3. "1P Coupled."*** Participants tapped along with an auditory metronome, while looking at the partner's hand tapping. Their own hand was video-recorded from above their shoulder, but swapping was performed so that the partner's hand was perceived from a 1st person perspective (as the partner would see oneself). Participants were explicitly asked to neglect the partner's movements and focus on following their own metronome. ***Condition 4. "2P Uncoupled" (control).*** Participants tapped along with an auditory metronome, while looking at their own hand. The partner's hand was video-recorded from a frontal position, but swapping was performed so that the own hand was perceived from a 2nd person perspective (as oneself would be seen by the partner).

and their standard errors for recurrence score, alongside the associated p values. In this analysis framework, parameter estimates provide a measure of effect size of straightforward interpretation for linear and non-linear changes over time, as long as the polynomial order is not too high.[59] With the interaction effect of Coupling and Perspective factors on the polynomial terms, we could quantify the specific effects of visual perspective on the evolution of the recurrence score when the partners were coupled.

Moving at the individual level of analysis, we tested whether the manipulation of visual perspectives induced significant changes in the experienced sense of ownership over the perceived hand. Aligned rank transform (ART) ANOVA[60] revealed significant main effects of Coupling (*Df residual* = 147, $F = 104.353$, $p < 0.001$) and Perspective (*Df residual* = 147, $F = 8.983$, $p < 0.01$) on the self-reported ownership ratings. The former indicates that participants were capable of telling apart their own hand from the partner's regardless of the visual perspective, whilst the latter indicates that perceiving a hand in 1st person generally resulted in a stronger sense of ownership. Crucially, the interaction effect between Coupling and Perspective (*Df residual* = 147, $F = 5.232$, $p < 0.05$) revealed that the increase in ownership relative to the 2nd person perspective was significantly stronger when participants were coupled. This means that the partner's hand, normally recognized as

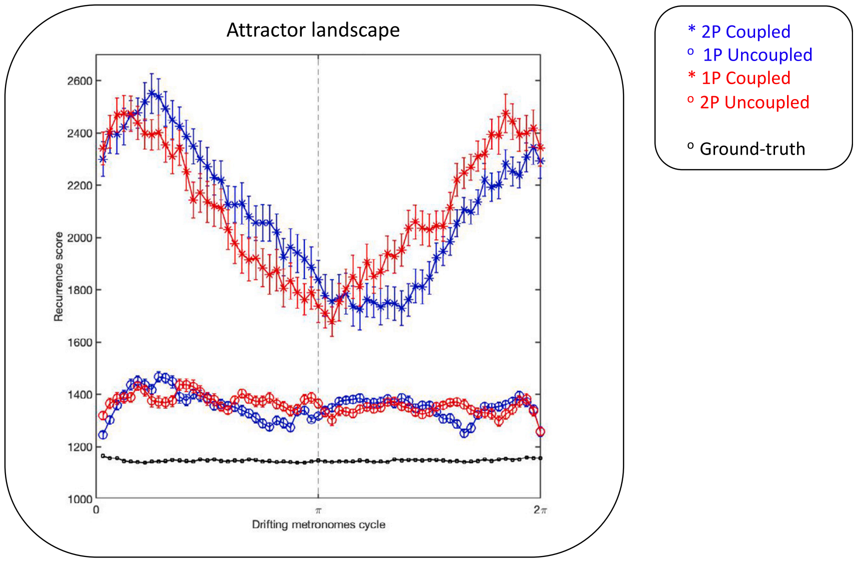


Figure 2. Attractor landscape

The time series depicted in the figure represent the evolution of the recurrence score as a function of the drifting metronomes' cycle, across experimental conditions. The grand-average was computed over the whole sample of dyads (N = 19), and for each dyad the time series was computed as the average of 10 consecutive cycles. Error bars indicate the standard error of the mean (SEM). For illustration and interpretation purposes, the black line in the plot shows the same analysis as performed on the two metronomes time series. This represents a ground-truth in the context of the paradigm, providing the reference recurrence score expected by a deterministic de-coupled system such as two linearly dephasing metronomes. A horizontal line lingering at the global minimum is the pattern expected from two partners when each of them is perfectly synchronizing with the assigned metronome, without influencing each other. The two time series just above the reference were computed from the uncoupled conditions (2 and 4), where each participant was tapping while seeing their own hand from 1st and 2nd person perspectives, respectively, so that no information was exchanged with the partner. As expected, the recurrence scores closely tracked the reference in these conditions, with random fluctuations around the mean and spurious recurrences due to human movement variability. Due to absence of coupling between the partners and hence their ignorance of the drifting metronomes' structure, it was feasible for them to follow the assigned metronomes. No significant difference was found between visual perspectives in uncoupled conditions. These control conditions provided a baseline for statistical contrasts, allowing us to assess the significance of eventual patterns deviating from the reference due to visual coupling. The two upper time series represent the coupled conditions (1 and 3), where each participant could see the hand of the partner from 2nd and 1st person perspectives, respectively. These are the critical conditions to focus on, in order to answer our main research question. When modeling empirical curves with orthogonal polynomials,[59] the intercepts of the fitted model were significantly greater than the uncoupled control conditions, capturing the main effect of Coupling. It is indeed clearly visible that both curves are on average above the respective controls. Focusing on the shape of the curves, it is also evident that both exhibit a parabolic curvature and a pronounced asymmetry. These two features were captured by the significant interaction effects of Coupling with the Quadratic and the Linear terms of Time, respectively.

Neither the average recurrence score nor the parabolic curvature significantly differed across levels of Perspective. However, we did find a significant interaction on the Linear term of Time, capturing a critical difference between the curves depending on visual perspective. Whilst in both coupled conditions the recurrence score reached a global minimum past the π midpoint, in 2nd person perspective it lingered on a longer horizontal trajectory into the second half of the cycle, before reaching the maximum with a steeper exponential growth. This resulted in a more prominent asymmetry which, as previously discussed,[6] indicates hysteresis in the system[2,9] since the rate of change of the recurrence score is dependent on the direction of the de-phasing (i.e., from 0 to π and from π to 0). The same in-phase attractor exerted a stronger "pull-back" on the dyad as it left the 0 point, followed by a steeper "push-forward" as it approached the same point at the end of the cycle. The anti-phase point can be seen as a "competition attractor",[9] for it facilitates de-coupling among the partners and pursuing of independent trajectories. This interpretation is empirically supported by our reference time series (black line in the plot), showing that horizontal line at a minimum occur when a de-phasing pattern is taking place. Crucially, the competition attractor around the π point resulted to be weaker when partners were coupled in 1st person perspective, since they did not manage to keep dephasing for quite as long. The dynamic balance shifted in favor of the cooperation attractor, resulting in a steeper increase of recurrence score. From these observations, we conclude that visual coupling in 1st person promotes the convergence of the dyadic system toward phase alignment as compared to the ecological mode of interaction in 2nd person.

belonging to somebody else, is perceived as belonging to one's own to a significantly greater extent due to the manipulation of visual perspectives. The same model was fit to the ratings of sense of ownership and sense of agency as experienced during the joint finger-tapping task. In this case, we only found significant main effect of Coupling on both ownership (*Df residual* = 147, *F* = 459.467, $p < 0.001$) and agency (*Df residual* = 147, *F* = 373.005, $p < 0.001$). This result indicates that in conditions of active movement, participants correctly attributed the hand and its actions to themselves, and did not experience illusory attribution of the partner's hand from any visual perspective. Median scores for self-reported ratings are shown in Figure 3.

Finally, the distributions of the relative phase between participants and the assigned metronome was computed for each condition, and the vector length R was calculated as a measure of individual performance. This measure quantifies the overall synchronization consistency during the whole trial. Rosso et al.[6] demonstrated that a reduced synchronization with the metronome in Coupled conditions, as contrasted to

Table 1. Recurrence score

	Recurrence score (N = 19)		
Predictors	*Estimate*	*SE*	*p*
(Intercept)	752.739	56.693	0.000
Time	−106.503	120.749	0.378
$Time^2$	4.933	127.602	0.969
Perspective	−0.180	73.777	0.998
Coupling	436.278 ***	73.777	<0.001
Time:Perspective	53.653	152.467	0.725
$Time^2$:Perspective	49.110	166.386	0.768
Time:Coupling	−122.145	152.467	0.423
$Time^2$:Coupling	1201.669 ***	166.386	<0.001
Perspective:Coupling	50.670	104.337	0.627
Time:Perspective:Coupling	478.850 *	215.621	0.026
$Time^2$:Perspective:Coupling	7.047	235.305	0. 976
	*$*p < 0.05$	**$p < 0.01$	***$p < 0.001$

Orthogonal polynomials model summary.

Uncoupled conditions, points to a systematic attraction toward the partner at the expense of complying with the task. The *perspective taking*, *empathic concern*, *fantasy* and *personal distress* scores of the *Interpersonal Reactivity Index* (IRI)[61] were included as continuous predictors in a factorial model (Coupling x Perspective) and fitted to the response variable R, to test whether personal empathic traits would modulate synchronization across conditions. None of the subscales of empathy had any significant effect on synchronization consistency. However, the model revealed a significant 2-way interaction between Coupling and Perspective (*Estimate* = −0.292, *SE* = 0.141, $p < 0.05$), indicating that the negative impact on the individual synchronization performance was stronger when the partner's hand was perceived from a 1st person perspective. This corroborates the results from the dyadic analyses, which showed that mutual attraction was enhanced in this condition. Figure 4 illustrates that being coupled to the partner leads to poorer performance in synchronizing with the metronome. Additionally, 1st person perspective improves the performance when participants see their own hand, but worsens it when seeing the partner's hand.

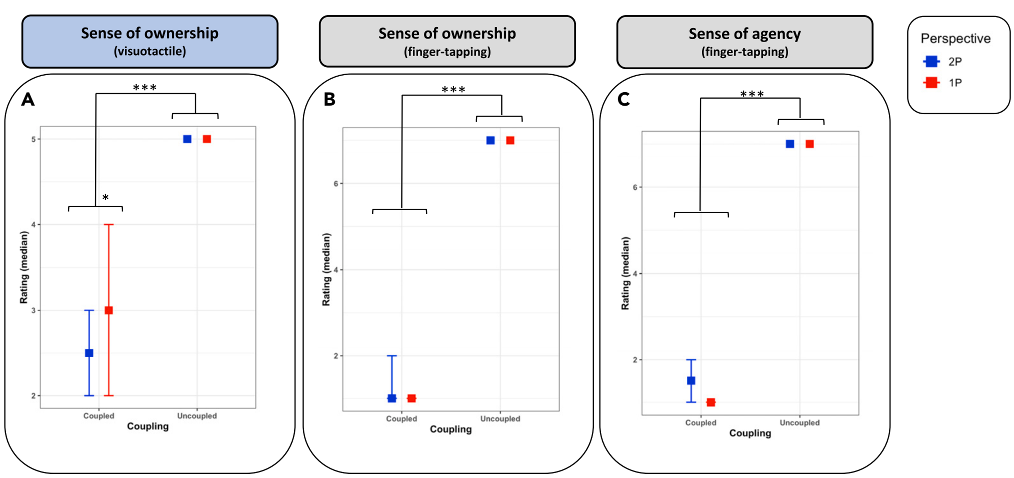


Figure 3. Sense of ownership over the perceived hand

Boxes represent median values of the subjective ratings referring to the following constructs (N = 37), across experimental conditions: (A) sense of ownership during visuotactile stimulation (scale 1–5), (B) sense of ownership during finger-tapping task (scale 1–7), (C) sense of agency during finger-tapping task (scale 1–7). Error bars represent the 95% confidence interval of the median. The fact that in Uncoupled conditions participants systematically recognized the hand as their own resulted in a ceiling effect, which did not allow to compute error bars in such conditions. For all response variables (A–C), we found a main effect of Coupling, whereas the interaction effect was significant only when sense of ownership was measured via visuotactile stimulation (A). Whilst the main effect is somewhat trivial, the interaction shows that the manipulation of visual perspectives was successful in inducing a subjective experience of embodiment over the partner's hand, specifically when this was perceived from a 1st person visual perspective. Asterisks represent the following p values: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

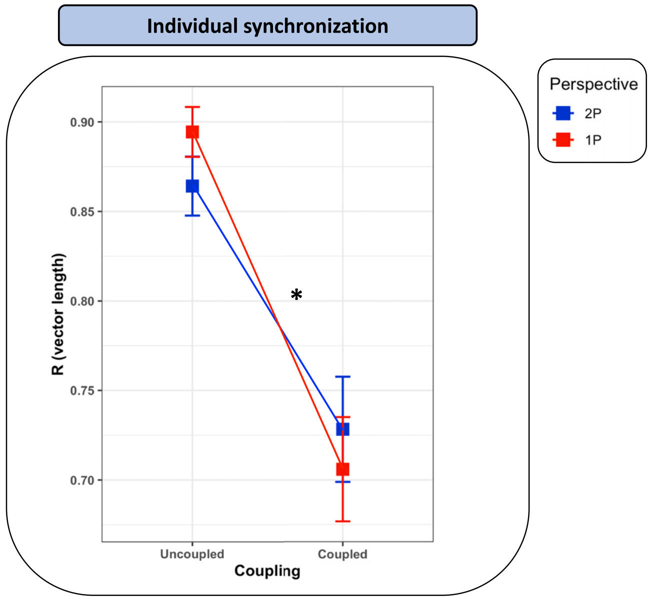


Figure 4. Individual synchronization consistency with assigned metronome

The dots represent the mean values of the vector length R of the relative phase distributions (N = 38), calculated as the difference between participants' and metronomes' phase time series. Error bars represent the standard error of the mean (SEM). The figure shows how synchronization performance tends to be poorer in coupled conditions because of the attraction toward the partner, which occurs at the expense of synchronizing with the assigned metronome. The significant interaction effect between Coupling and Perspective is highlighted: 1st person perspective improves the performance when seeing one's own hand, while it makes it worse when seeing the partner's hand. We highlight that the average performance in uncoupled conditions was optimal, as indicated by R values close to 1. Asterisks represent the following p values: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

The same models fitted on the subjective ratings of embodiment revealed a significant 2-way interaction effect between *empathic concern* and Coupling (*Estimate* = −0.177, *SE* = 0. 089, p = 0.05). This indicates that participants with higher empathic concern experienced a stronger sense of ownership over the partner's hand. We also point out a trend toward 3-way interaction with Coupling and Perspective (*Estimate* = 0.207, *SE* = 0.128, p = 0.11), showing that the effect tended to be stronger when the partner's hand was perceived in 1st person. Models' summaries for the Individual level of analysis are reported in Tables 2, 3, and 4. In all models, Uncoupled (factor Coupling) and 2P (factor Perspective) were set as 0-levels for statistical contrasts.

DISCUSSION

The present study investigated the role of visual perspective in spontaneous dyadic entrainment, which is considered to be the most minimal and fundamental level of rhythmic interpersonal coordination.[9,62,63] By inducing a hand-swap illusion between partners engaged in joint finger-tapping, we were able to quantify overall synchronization strength and local attractor dynamics when they could perceive each other's hand in 1st person, and compare them to an ecological mode of interaction in 2nd person. The drifting metronomes paradigm was adopted to guide the partners through a systematic exploration of their attractor landscape, to detect attractor points over the whole space of coordinative states.[6] As analysis framework, joint recurrence quantification analysis (JRQA)[58] yielded a relational measure to quantify the degree of temporal coordination within the dyad throughout the task.[9]

In the first place, we were able to replicate the results from our previous report.[6] When participants were visually coupled in a 2nd person face-to-face interaction, a cooperation process dominated the interaction, resulting in recurrent states of coordinated behavior despite the active attempt of neglecting the partner's rhythm and pursue individual trajectories. Crucially, the effect was not constant over the whole drifting metronomes' cycle, but rather modulated by consistent attractor dynamics. As dyads were driven by the metronomes through the space of relative phase values, we could observe the recurrence score oscillating between global maxima and global minima in proximity of critical regions. High recurrence score indicates high degree of temporal coordination within the dyad, while low recurrence score indicates temporal independence. When the recurrence score lingers at baseline levels for a sustained period of time, it means that partners managed to ignore each other and maintain their own tempo, tracking the de-phasing pattern of the drifting metronomes. It is in the transitions over these critical regions that the dynamic balance between two opposite tendencies of the system can be observed, namely the "pull" into temporally coordinated behavior and the "push" toward de-coupled, independent behavior. The maxima and minima of the recurrence score were observed around the in-phase (0) and anti-phase (π) points, which operated as "cooperation attractor" and "competition attractor,"[2] respectively. Whilst the partners tended to move together at a collective level of coupled behavior in proximity of the cooperation attractor, it became easier for them to pursue independent de-coupled trajectories in proximity of the competition attractor.

Table 2. Sense of ownership

	Sense of ownership – visuotactile stimulation (N = 37)		
Predictors	*Df residual*	*F value*	*p*
Coupling	147	104.353***	<0.001
Perspective	147	8.983**	0.003
Coupling:Perspective	147	5.232*	0.023
	**p* < 0.05	***p* < 0.01	****p* < 0.001

	Sense of ownership – finger-tapping task (N = 37)		
Predictors	*Df residual*	*F value*	*p*
Coupling	147	459.467***	<0.001
Perspective	147	2.356	0.127
Coupling:Perspective	147	1.278	0.260
	**p* < 0.05	***p* < 0.01	****p* < 0.001

	Sense of agency – finger-tapping task (N = 37)		
Predictors	*Df residual*	*F value*	*p*
Coupling	147	373.005***	<0.001
Perspective	147	0.063	0.802
Coupling:Perspective	147	0.021	0.886
	**p* < 0.05	***p* < 0.01	****p* < 0.001

ART 2-way ANOVA models summaries.

Moving on to the present experimental design, the main effect of Coupling revealed that recurrence score was on average significantly higher in coupled conditions (global cooperation), while its 2-way interaction with the quadratic component of Time showed a significant modulation by attractor points over the course of the drifting metronomes' cycle (local dynamics). More specifically, recurrence score grew as a parabolic function of the relative phase between metronomes, finding its maximum around the in-phase point and its minimum right after the middle anti-phase point. As shown in Figure 2, when participants were visually coupled, both 1st and 2nd person perspectives scored on average above the baseline levels of the uncoupled control conditions, and exhibited the same depth of the parabolic curvature. Crucially, although the mutual assumption of 1st person perspective did not affect the global level of recurrence, it resulted in a stronger attraction toward the cooperation attractor in the second half of the cycle. The effect was captured by the significant 3-way interaction between Coupling, Perspective, and the linear term of Time, which indicates that the asymmetry in the parabolic curve significantly changes depending on the levels of perspective. Due to the manipulation, the pull toward the in-phase attractor began earlier on in the cycle, such that participants did not manage to take advantage of the π region to de-couple. In our paradigm, the relative phase between the metronomes was manipulated as a control parameter[64] from 0 to π (ascending) and from π to 0 (descending) radians. In this scenario, dyads

Table 3. Vector length R (synchronization with metronomes)

	Vector length R (N = 38)		
Predictors	*Estimate*	*SE*	*p*
(Intercept)	0.815	0.101	0.000
EC	0.003	0.007	0.618
Coupling	−0.139	0.098	0.160
Perspective	−0.038	0.098	0.694
Coupling:Perspective	−0.292*	0.140	0.041
EC:Coupling	0.002	0.007	0.697
EC:Perspective	0.005	0.007	0.486
EC:Coupling:Perspective	0.016	0.009	0.100
	**p* < 0.05	***p* < 0.01	****p* < 0.001

Mixed-effects model summary. "Empathic concern" is abbreviated as "EC."

Table 4. Sense of ownership

	Sense of ownership (N = 37)		
Predictors	*Estimate*	*SE*	*p*
(Intercept)	4.768	1.016	0.000
EC	−0.013	0.068	0.850
Coupling	0.643	1.333	0.6306
Perspective	−0.184	1.333	0.890
Coupling:Perspective	−2.922	1.908	0.129
EC:Coupling	−0.177	0.089	0.050
EC:Perspective	0.025	0.089	0.779
EC:Coupling:Perspective	0.207	0.128	0.110
	$*p < 0.05$	*$**p < 0.01$*	*$***p < 0.001$*

Mixed-effects model summary. "Empathic concern" is abbreviated as "EC."

tended to slowly transition from cooperation to competition as the metronomes' relative phase diverged from 0 to π, whereas they exhibited a more abrupt transition from competition to cooperation as metronomes converged from π to 0. The asymmetry, observed in the condition of 2nd person and captured by the linear term of Time in the polynomial model, was previously reported in the paradigm[6] as the manifestation of hysteresis, namely the dependency of the dyadic system on history and directionality of the interaction.[2] We therefore interpret a steeper growth of the recurrence curves as a sign of increased hysteresis and stronger cooperation attractor when participants were coupled in 1st person perspective.[9]

When explaining our finding, it is important to first consider whether differences in coupling strength could be explained by varying amounts of information in the visual percept. Apart from being rotated by 180°, detailed features of the observed hand such as geometry, color, and texture, were held constant across conditions (see Figure 1). However, in order to maintain the natural proportions of the human body, the hand occupied a wider angle of the visual scene when it was presented in 1st person perspective. Although this variation might have enhanced perceptual salience, it was a necessary compromise in order to achieve a realistic percept, since the participants' own hand was effectively closer to their eyes as compared to the hand of the partner sitting across the table. As the change in perspective is the critical variable, our explanation should focus on how the brain processes bodily effectors within an egocentric frame of reference and explain how embodying an effector facilitates motor adaptation to temporal mismatches between performed and observed actions, leading to stronger entrainment as an emergent property of dyadic behavior.

We propose two possible explanations for this phenomenon. The first one, based on social cognition, frames interpersonal synchronization in terms of self-other integration,[34,65–73] where distinct but overlapping brain networks[67,74–76] process information related to one's own movements and those observed in the other. To account for the dynamic development of dyadic rhythmic interactions, it was recently proposed that self-other integration and segregation are two metastable attractor states underlying coupled and de-coupled behaviors, respectively.[66] Such metastability between integration/segregation states maps well onto the cooperation/competition attractors we discussed so far. According to this view, transitions occur as the brain selects whether it is more efficient to integrate perceptual information in one merged model for self and other, or hold two segregated models to attribute the perceived action to one of the two agents. The visuospatial overlap between the other's hand perceived in 1st person and one's own hand would blur the difference between self and other, promoting the merging of two separate models into one, tightening action-perception loops between the partners, and ultimately strengthening the attraction toward a coupled state.

The second explanation is based on a purely sensorimotor account, it does not call into question the representation of the other, and is based on the fact that our manipulation set the conditions for the brain to represent the perceived effector as actually belonging to the bodily self.[19–23,26,40,77] The mapping between actions and their sensory consequences is learned throughout a lifetime of sensorimotor contingencies and action-perception dependencies.[78] By direct experience, everybody learns to expect a temporal match between a movement and its visual feedback as perceived in 1st person. This suggests that the brain employs forward models[41,42] for actions generated by the embodied hand, and would in turn engage in error correction when the observed movement does not temporally match the predicted outcome of the executed movement. Sensory prediction errors drive motor adaptation in terms of movement trajectory,[79] velocity,[80,81] and timing.[44,82] When error minimization is carried out simultaneously by two coupled individuals with respect to each other's movements, the process ultimately leads to dyadic synchronized behavior.[34,36,83] Interestingly, the mirror neurons system[84] literature suggests that activating inverse and forward models via action observation requires a visuospatial transformation process, to remap the movement into an egocentric frame of reference.[11] This supports the idea that spatial perspective-taking is an embodied cognitive process, in fact the internal emulation of a physical alignment of perspectives.[12] Noteworthy, such transformation comes with a processing cost which is a function of the angular disparity between the observer and the actor,[12–17] as if the observer was internally simulating a rotation into the other's point of view. In a dyadic setting, such putative process can be bypassed with the technological means deployed in the present study.

We point out that the sensorimotor account is more parsimonious compared to the socio-cognitive one, and more conservative in its assumptions. Whereas general theories of brain functioning differ in the assumption that individuals mutually adapt their behaviors based on internal models of the other, they are unified by the shared principle of error minimization.[85] The validation of either theory is out of the scope of the present work, and we are far from providing conclusive evidence on this matter. Therefore, we lean in favor of the more parsimonious interpretation based on error correction as tenet shared principle. We highlight that our manipulation effectively strengthened the coupling between individuals by inducing the embodiment of the partner's hand, and propose that embodiment engaged error correction mechanisms to maintain the temporal match between executed and perceived movements. Dyadic entrainment emerges here as property of collective behavior from the interaction of two individuals mutually adjusting for their respective prediction errors.

To complement the behavioral findings, we hereby discuss the subjective experiential correlates of mutual embodiment via body-swap. During the visuo-tactile stimulation procedure, participants were capable of systematically discriminating their own hand from the partner's, and showed a general preference for either hand when viewed from a 1st person perspective. These results came with no surprise, since behavioral and neurophysiological evidence supports the existence of a bodily self recognition mechanism reliant on visual and sensorimotor representations of the hand,[86] while the hand orientation with respect to bodily coordinates is a crucial aspect to meet the conditions for embodiment.[27] The crucial finding is the interaction effect between factors, showing that the hand of a partner can in fact be integrated in the bodily representation of the self to a greater extent when visually presented in 1st person, as compared to the 2nd person perspective (see Figure 3A). These results confirm the success of our body-swap procedure in eliciting an experiential counterpart to our behavioral findings, suggesting that the putative overlap of self-other representations thought to underpin interpersonal synchronization[34,66] may leverage on the plasticity of bodily representations.[19,20] In sum, the explicit measures based on self-reports are coherent with the implicit measures based on attractor dynamic within the dyad. Nevertheless, when asked about sense of ownership and sense of agency experienced during the joint finger-tapping task, participants could very well discriminate between their own's and their partner's hands regardless of the visual perspective (see Figures 3B and 3C). Sense of ownership subsists as long as sensorimotor congruency is maintained,[25,33,87] and sense of agency breaks down when the timing of sensory feedback does not match the prediction.[25,88] However, temporal congruency was not constant throughout the task, but rather varying as a function of the drifting metronomes' cycles, which likely led to the breakdown of both illusions of ownership and agency. We should point out that questionnaire items suffer the major flaw of referring to the task as a whole, whereas from our standpoint the most interesting behavioral findings came from an analysis of local dynamics over the course of the interaction.

Finally, we did not find evidence for any association between empathic traits and the proneness to synchronize with the partner when coupled in any visual perspective. This was unexpected, since the link appears to be well documented in the literature[46] and resonates with the idea that attraction to coordinated states is to some extent informative of the most minimal socioemotional connectedness.[9] We propose that the negative finding may be attributed to the rigorous competitive nature of the drifting metronomes. Arguably, when task constraints are looser, there is more margin for more empathic participants to intentionally cooperate with each other. On the other hand, the observation that recurrence score systematically dropped as the partners got further away from the cooperation attractor region, shows that our participants consistently attempted to comply with the instruction to intentionally neglect the partner. To the best of our knowledge, no study has previously investigated the contribution of empathy to spontaneous interpersonal synchronization in a comparable experimental situation. We conclude that empathy is not a significant predictor when the intention to synchronize is constrained by the task, which in this case revealed a dissociation between low-level dyadic entrainment and high-level cognitive empathy. The negative finding suggests that intentionality may act as a mediating variable in some of the numerous studies reporting significant correlations between empathy and interpersonal synchrony. However, among the dimensions of empathy considered here,[61] empathic concern stood out as predictor for the subjective ratings of embodiment. Specifically, higher emotionality and concern for others predicted a stronger inclination to experience ownership over the hand of another person. This points at the interplay between empathy and the mechanisms underlying embodiment, transferring evidence for such associations from the VR literature[48–51] to partial body-swap with a real human partner. This particular dimension of trait empathy should not be neglected when adopting the technology for real-world applications, since the evidence suggests it may be a personal variable relevant to the outcome of the manipulation.

Conclusions

The main aim of the present work was to assess whether interpersonal synchronization can be facilitated by experimentally inducing the 1st person view of a partner during a rhythmic interaction, as compared to the ecological mode of interaction in 2nd person. Our results support the idea that such manipulation strengthens the coupling between interacting individuals, promoting the cooperation process which facilitates the units of a dyadic system to move together at the collective level of behavior.[6] From a socio-cognitive viewpoint, we put forward that the dynamic balance between cooperation and competition processes may be underpinned by metastable self-other integration and segregation processes taking place in individual brains during the interaction.[66] The induction of mutual embodiment would then facilitate transitions toward integration and cooperation within the dyad. Whilst this interpretation is plausible, the study does not provide conclusive evidence for a socio-cognitive account. We argue that an explanation based on sensory prediction and adaptation in motor control[78] would be more conservative, for it does not make assumptions on the representation of the other in the brain, while accounting for error correction mechanisms leading to dyadic entrainment as emergent property of the interaction.

The major fundamental contribution of our work lies in the observation that dyadic coordination dynamics are subject to the manipulation of visual perspective. Whatever the cognitive mechanism behind, our findings show that this manipulation can be used to steer social interactions, supporting joint action by enhancing interpersonal synchronization. Based on our findings, we propose that a technology informed by principles of body-swapping[33] holds considerable potential for facilitating interpersonal coordination in a variety of contexts, including motor training, sports, music education, and rehabilitation.

Limitations of the study

The major limitation of our study is that, although it provides a descriptive account of the effects of visual perspective on interpersonal synchronization, it does not systematically test the underlying explanatory mechanisms through computational models. We suggest that future work is needed to systematically evaluate the explanatory capabilities of the models referenced in our discussion.

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Pre-processing
  - Phase-space reconstruction
  - Statistical models

ACKNOWLEDGMENTS

The present study was funded by Bijzonder Onderzoeksfonds (BOF) from Ghent University (Belgium), in the context of a joint-PhD project with the University of Lille (France) (I-SITE ULNE program; grant number: 01D21819). The authors are grateful to Ivan Schepers for building the hardware of the finger-tapping device, to Kevin Smink for the illustration of the experimental design, and to Canan Nuran Gener for her precious help in collecting the data.

AUTHOR CONTRIBUTIONS

Conceptualization, M.R.; Methodology, M.R. and B.v.K.; Software, M.R. and B.v.K.; Validation, M.R. and B.v.K.; Formal analysis, M.R.; Investigation, M.R. and B.v.K.; Resources, M.R. and B.v.K.; Data Curation, M.R. and B.v.K.; Writing – Original Draft, M.R.; Writing – Review and Editing, B.v.K., P.M., M.L.; Visualization, M.R.; Supervision, P.M. and M.L.; Project Administration, M.R.; Funding Acquisition, M.R. and M.L.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 4, 2023
Revised: June 20, 2023
Accepted: September 27, 2023
Published: September 29, 2023

REFERENCES

1. Richardson, M.J., Marsh, K.L., and Schmidt, R.C. (2005). Effects of visual and verbal interaction on unintentional interpersonal coordination. J. Exp. Psychol. Hum. Percept. Perform. *31*, 62–79.
2. Richardson, M.J., Marsh, K.L., and Baron, R.M. (2007). Judging and actualizing intrapersonal and interpersonal affordances. J. Exp. Psychol. Hum. Percept. Perform. *33*, 845–859.
3. Crombé, K., Denys, M., and Maes, P.-J. (2022). The role of a mechanical coupling in (spontaneous) interpersonal synchronization: A human version of Huygens' clock experiments. In Timing Time Percept, pp. 1–20.
4. Schmidt, R.C., and O'Brien, B. (1997). Evaluating the Dynamics of Unintended Interpersonal Coordination. Ecol. Psychol. *9*, 189–206.
5. Schmidt, R.C., Carello, C., and Turvey, M.T. (1990). Phase transitions and critical fluctuations in the visual coordination of rhythmic movements between people. J. Exp. Psychol. Hum. Percept. Perform. *16*, 227–247.
6. Rosso, M., Maes, P.J., and Leman, M. (2021). Modality-specific attractor dynamics in dyadic entrainment. Sci. Rep. *11*, 18355.
7. Kelso, J.A.S. (1995). Dynamic Patterns: The Self-Organization of Brain and Behavior (MIT Press).
8. Tognoli, E., Zhang, M., Fuchs, A., Beetle, C., and Kelso, J.A.S. (2020). Coordination Dynamics: A Foundation for Understanding Social Behavior. Front. Hum. Neurosci. *14*, 317.
9. Marsh, K.L., Richardson, M.J., and Schmidt, R.C. (2009). Social connection through joint action and interpersonal coordination. Top. Cogn. Sci. *1*, 320–339.
10. Gallese, V., and Sinigaglia, C. (2011). What is so special about embodied simulation? Trends Cogn. Sci. *15*, 512–519.
11. Oh, H., Braun, A.R., Reggia, J.A., and Gentili, R.J. (2019). Fronto-parietal mirror neuron system modeling: Visuospatial transformations support imitation learning

independently of imitator perspective. Hum. Mov. Sci. 65, 121–141. https://doi.org/10.1016/j.humov.2018.05.013.
12. Kessler, K., and Thomson, L.A. (2010). The embodied nature of spatial perspective taking: embodied transformation versus sensorimotor interference. Cognition 114, 72–88.
13. Kessler, K. (2000). Spatial Cognition and Verbal Localisations: A Connectionist Model for the Interpretation of Spatial Prepositions (Deutscher Universitäts-Verlag).
14. May, M. (2004). Imaginal perspective switches in remembered environments: transformation versus interference accounts. Cogn. Psychol. 48, 163–206.
15. Wraga, M., Shephard, J.M., Church, J.A., Inati, S., and Kosslyn, S.M. (2005). Imagined rotations of self versus objects: an fMRI study. Neuropsychologia 43, 1351–1361.
16. Zacks, J.M., and Michelon, P. (2005). Transformations of visuospatial images. Behav. Cogn. Neurosci. Rev. 4, 96–118.
17. Keehner, M., Guerin, S.A., Miller, M.B., Turk, D.J., and Hegarty, M. (2006). Modulation of neural activity by angle of rotation during imagined spatial transformations. Neuroimage 33, 391–398.
18. Petkova, V.I., Björnsdotter, M., Gentile, G., Jonsson, T., Li, T.-Q., and Ehrsson, H.H. (2011). From part- to whole-body ownership in the multisensory brain. Curr. Biol. 21, 1118–1122.
19. Tsakiris, M. (2010). My body in the brain: a neurocognitive model of body-ownership. Neuropsychologia 48, 703–712.
20. Tsakiris, M. (2017). The multisensory basis of the self: From body to identity to others [Formula: see text]. Q. J. Exp. Psychol. 70, 597–609.
21. Botvinick, M., and Cohen, J. (1998). Rubber hands "feel" touch that eyes see. Nature 391, 756.
22. Slater, M., Perez-Marcos, D., Ehrsson, H.H., and Sanchez-Vives, M.V. (2008). Towards a digital body: the virtual arm illusion. Front. Hum. Neurosci. 2, 6.
23. Slater, M., Perez-Marcos, D., Ehrsson, H.H., and Sanchez-Vives, M.V. (2009). Inducing illusory ownership of a virtual body. Front. Neurosci. 3, 214–220.
24. Kilteni, K., Groten, R., and Slater, M. (2012). The sense of embodiment in virtual reality. In Presence: Teleoperators and Virtual.
25. Kalckert, A., and Ehrsson, H.H. (2012). Moving a Rubber Hand that Feels Like Your Own: A Dissociation of Ownership and Agency. Front. Hum. Neurosci. 6, 40.
26. Kalckert, A., and Ehrsson, H.H. (2014). The moving rubber hand illusion revisited: comparing movements and visuotactile stimulation to induce illusory ownership. Conscious. Cogn. 26, 117–132.
27. 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.
28. Maselli, A., and Slater, M. (2013). The building blocks of the full body ownership illusion. Front. Hum. Neurosci. 7, 83.
29. Petkova, V.I., Khoshnevis, M., and Ehrsson, H.H. (2011). The perspective matters! Multisensory integration in ego-centric reference frames determines full-body ownership. Front. Psychol. 2, 35.
30. Serino, A., Alsmith, A., Costantini, M., Mandrigin, A., Tajadura-Jimenez, A., and Lopez, C. (2013). Bodily ownership and self-location: components of bodily self-consciousness. Conscious. Cogn 22, 1239–1252.
31. Ehrsson, H.H. (2007). The experimental induction of out-of-body experiences. Science 317, 1048.
32. Guterstam, A., and Ehrsson, H.H. (2012). Disowning one's seen real body during an out-of-body illusion. Conscious. Cogn. 21, 1037–1042.
33. Petkova, V.I., and Ehrsson, H.H. (2008). If I were you: perceptual illusion of body swapping. PLoS One 3, e3832.
34. Koban, L., Ramamoorthy, A., and Konvalinka, I. (2019). Why do we fall into sync with others? Interpersonal synchronization and the brain's optimization principle. Soc. Neurosci. 14, 1–9.
35. Clark, A. (1999). Where brain, body, and world collide. Cogn. Syst. Res. 1, 5–17.
36. Konvalinka, I., Vuust, P., Roepstorff, A., and Frith, C.D. (2010). Follow you, follow me: continuous mutual prediction and adaptation in joint tapping. Q. J. Exp. Psychol. 63, 2220–2230.
37. Heggli, O.A., Konvalinka, I., Kringelbach, M.L., and Vuust, P. (2019). Musical interaction is influenced by underlying predictive models and musical expertise. Sci. Rep. 9, 11048.
38. Dell'Anna, A., Fossataro, C., Burin, D., Bruno, V., Salatino, A., Garbarini, F., Pia, L., Ricci, R., Leman, M., and Berti, A. (2018). Entrainment beyond embodiment. Neuropsychologia 119, 233–240.
39. Gallese, V., and Sinigaglia, C. (2010). The bodily self as power for action. Neuropsychologia 48, 746–755.
40. Della Gatta, F., Garbarini, F., Puglisi, G., Leonetti, A., Berti, A., and Borroni, P. (2016). Decreased motor cortex excitability mirrors own hand disembodiment during the rubber hand illusion. Elife 5, e14972. https://doi.org/10.7554/eLife.14972.
41. Blakemore, S.J., Frith, C.D., and Wolpert, D.M. (1999). Spatio-temporal prediction modulates the perception of self-produced stimuli. J. Cogn. Neurosci. 11, 551–559.
42. Blakemore, S.-J. (2017). Why can't you tickle yourself? In The Anatomy of Laughter (Routledge), pp. 34–41.
43. Wolpert, D.M., Doya, K., and Kawato, M. (2003). A unifying computational framework for motor control and social interaction. Philos. Trans. R. Soc. Lond. B Biol. Sci. 358, 593–602.
44. Maes, P.-J., Giacofci, M., and Leman, M. (2015). Auditory and motor contributions to the timing of melodies under cognitive load. J. Exp. Psychol. Hum. Percept. Perform. 41, 1336–1352.
45. Tuller, B., and Kelso, J.A. (1989). Environmentally-specified patterns of movement coordination in normal and split-brain subjects. Exp. Brain Res. 75, 306–316.
46. Tzanaki, P. (2022). The Positive Feedback Loop of Empathy and Interpersonal Synchronisation: Discussing a Theoretical Model and its Implications for Musical and Social Development. Music Sci. 5. 20592043221142716.
47. Maister, L., Slater, M., Sanchez-Vives, M.V., and Tsakiris, M. (2015). Changing bodies changes minds: owning another body affects social cognition. Trends Cogn. Sci. 19, 6–12.
48. Peck, T.C., Seinfeld, S., Aglioti, S.M., and Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. Conscious. Cogn. 22, 779–787.
49. Thériault, R., Olson, J.A., Krol, S.A., and Raz, A. (2021). Body swapping with a Black person boosts empathy: Using virtual reality to embody another. Q. J. Exp. Psychol. 74, 2057–2074.
50. Peck, T.C., Doan, M., Bourne, K.A., and Good, J.J. (2018). The Effect of Gender Body-Swap Illusions on Working Memory and Stereotype Threat. IEEE Trans. Vis. Comput. Graph. 24, 1604–1612.
51. Peck, T.C., Good, J.J., and Bourne, K.A. (2020). Inducing and Mitigating Stereotype Threat Through Gendered Virtual Body-Swap Illusions. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (Association for Computing Machinery), pp. 1–13.
52. Paladino, M.-P., Mazzurega, M., Pavani, F., and Schubert, T.W. (2010). Synchronous multisensory stimulation blurs self-other boundaries. Psychol. Sci. 21, 1202–1207.
53. Tajadura-Jiménez, A., Grehl, S., and Tsakiris, M. (2012). The other in me: interpersonal multisensory stimulation changes the mental representation of the self. PLoS One 7, e40682.
54. Ventre-Dominey, J., Gibert, G., Bosse-Platiere, M., Farnè, A., Dominey, P.F., and Pavani, F. (2019). Embodiment into a robot increases its acceptability. Sci. Rep. 9, 10083.
55. Gallese, V. (2019). Embodied simulation. Its bearing on aesthetic experience and the dialogue between neuroscience and the humanities. Gestalt Theory 41, 113–127.
56. Novembre, G., Ticini, L.F., Schütz-Bosbach, S., and Keller, P.E. (2012). Distinguishing self and other in joint action. Evidence from a musical paradigm. Cereb. Cortex 22, 2894–2903.
57. Novembre, G., Mitsopoulos, Z., and Keller, P.E. (2019). Empathic perspective taking promotes interpersonal coordination through music. Sci. Rep. 9, 12255.
58. Marwan, N., Carmen Romano, M., Thiel, M., and Kurths, J. (2007). Recurrence plots for the analysis of complex systems. Phys. Rep. 438, 237–329.
59. Mirman, D. (2017). Growth Curve Analysis and Visualization Using R (CRC Press).
60. Wobbrock, J.O., Findlater, L., Gergle, D., and Higgins, J.J. (2011). The aligned rank transform for nonparametric factorial analyses using only anova procedures. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI '11 (Association for Computing Machinery), pp. 143–146.
61. Davis, M.H.; Others (1980). A multidimensional approach to individual differences in empathy. JSAS Catalog of Selected Documents in Psychology 10, 85.
62. Sebanz, N., and Knoblich, G. (2009). Prediction in joint action: what, when, and where. Top. Cogn. Sci. 1, 353–367.
63. Knoblich, G., and Sebanz, N. (2008). Evolving intentions for social interaction: from entrainment to joint action. Philos. Trans. R. Soc. Lond. B Biol. Sci. 363, 2021–2031.
64. Strogatz, S., Friedman, M., Mallinckrodt, A.J., and McKay, S. (1994). Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering. Comput. Phys. Commun. 8, 532.
65. Heggli, O.A., Cabral, J., Konvalinka, I., Vuust, P., and Kringelbach, M.L. (2019). A Kuramoto model of self-other integration

across interpersonal synchronization strategies. PLoS Comput. Biol. *15*, e1007422.
66. Heggli, O.A., Konvalinka, I., Kringelbach, M.L., and Vuust, P. (2021). A metastable attractor model of self–other integration (MEAMSO) in rhythmic synchronization. Philos. Trans. R. Soc. Lond. B Biol. Sci. *376*, 20200332.
67. Farrer, C., and Frith, C.D. (2002). Experiencing Oneself vs Another Person as Being the Cause of an Action: The Neural Correlates of the Experience of Agency. Neuroimage *15*, 596–603.
68. Blakemore, S.J., and Frith, C. (2003). Self-awareness and action. Curr. Opin. Neurobiol. *13*, 219–224.
69. van der Meer, L., Groenewold, N.A., Nolen, W.A., Pijnenborg, M., and Aleman, A. (2011). Inhibit yourself and understand the other: neural basis of distinct processes underlying Theory of Mind. Neuroimage *56*, 2364–2374.
70. Novembre, G., Sammler, D., and Keller, P.E. (2016). Neural alpha oscillations index the balance between self-other integration and segregation in real-time joint action. Neuropsychologia *89*, 414–425.
71. Huberth, M., Dauer, T., Nanou, C., Román, I., Gang, N., Reid, W., Wright, M., and Fujioka, T. (2019). Performance monitoring of self and other in a turn-taking piano duet: A dual-EEG study. Soc. Neurosci. *14*, 449–461.
72. Varlet, M., Nozaradan, S., Nijhuis, P., and Keller, P.E. (2020). Neural tracking and integration of "self" and "other" in improvised interpersonal coordination. Neuroimage *206*, 116303.
73. Dumas, G., Moreau, Q., Tognoli, E., and Kelso, J.A.S. (2020). The Human Dynamic Clamp Reveals the Fronto-Parietal Network Linking Real-Time Social Coordination and Cognition. Cereb. Cortex *30*, 3271–3285.
74. Mitchell, J.P., Banaji, M.R., and Macrae, C.N. (2005). The link between social cognition and self-referential thought in the medial prefrontal cortex. J. Cogn. Neurosci. *17*, 1306–1315.
75. Hasson, U., and Frith, C.D. (2016). Mirroring and beyond: coupled dynamics as a generalized framework for modelling social interactions. Philos. Trans. R. Soc. Lond. B Biol. Sci. *371*, 20150366. https://doi.org/10.1098/rstb.2015.0366.
76. Frith, C.D., and Frith, U. (1999). Interacting minds–a biological basis. Science *286*, 1692–1695.
77. Tsakiris, M., and Haggard, P. (2005). The rubber hand illusion revisited: visuotactile integration and self-attribution. J. Exp. Psychol. Hum. Percept. Perform. *31*, 80–91.
78. Shadmehr, R., Smith, M.A., and Krakauer, J.W. (2010). Error correction, sensory prediction, and adaptation in motor control. Annu. Rev. Neurosci. *33*, 89–108.
79. Mazzoni, P., and Krakauer, J.W. (2006). An implicit plan overrides an explicit strategy during visuomotor adaptation. J. Neurosci. *26*, 3642–3645.
80. Smith, M.A., Ghazizadeh, A., and Shadmehr, R. (2006). Interacting adaptive processes with different timescales underlie short-term motor learning. PLoS Biol. *4*, e179.
81. Wagner, M.J., and Smith, M.A. (2008). Shared internal models for feedforward and feedback control. J. Neurosci. *28*, 10663–10673.
82. Furuya, S., and Soechting, J.F. (2010). Role of auditory feedback in the control of successive keystrokes during piano playing. Exp. Brain Res. *204*, 223–237.
83. Rosso, M., Heggli, O.A., Maes, P.J., Vuust, P., and Leman, M. (2022). Mutual beta power modulation in dyadic entrainment. Neuroimage *257*, 119326.
84. Heyes, C., and Catmur, C. (2022). What Happened to Mirror Neurons? Perspect. Psychol. Sci. *17*, 153–168.
85. Palmer, C., and Demos, A.P. (2021). Are we in time? How predictive coding and dynamical systems explain musical synchrony. Curr. Dir. Psychol. Sci. *31*, 147–153.
86. Galigani, M., Ronga, I., Fossataro, C., Bruno, V., Castellani, N., Rossi Sebastiano, A., Forster, B., and Garbarini, F. (2021). Like the back of my hand: Visual ERPs reveal a specific change detection mechanism for the bodily self. Cortex *134*, 239–252.
87. Kilteni, K., Maselli, A., Kording, K.P., and Slater, M. (2015). Over my fake body: body ownership illusions for studying the multisensory basis of own-body perception. Front. Hum. Neurosci. *9*, 141.
88. Haggard, P., Clark, S., and Kalogeras, J. (2002). Voluntary action and conscious awareness. Nat. Neurosci. *5*, 382–385.
89. Oldfield, R.C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia *9*, 97–113.
90. Rosso, M., Leman, M., and Moumdjian, L. (2021). Neural Entrainment Meets Behavior: The Stability Index as a Neural Outcome Measure of Auditory-Motor Coupling. Front. Hum. Neurosci. *15*, 668918.
91. Takens, F. (1981). Detecting strange attractors in turbulence. In Dynamical Systems and Turbulence, Warwick 1980 (Springer Berlin Heidelberg), pp. 366–381.
92. Afsar, O., Tirnakli, U., and Marwan, N. (2018). Recurrence Quantification Analysis at work: Quasi-periodicity based interpretation of gait force profiles for patients with Parkinson disease. Sci. Rep. *8*, 9102.
93. Fraser, A.M., and Swinney, H.L. (1986). Independent coordinates for strange attractors from mutual information. Phys. Rev. A Gen. Phys. *33*, 1134–1140.
94. Bradley, E., and Kantz, H. (2015). Nonlinear time-series analysis revisited. Chaos *25*, 097610.
95. Rulkov, N.F., Sushchik, M.M., Tsimring, L.S., and Abarbanel, H.D. (1995). Generalized synchronization of chaos in directionally coupled chaotic systems. Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics *51*, 980–994.
96. Eckmann, J.-P., Kamphorst, S.O., and Ruelle, D. (1995). Recurrence plots of dynamical systems. World Sci. Series Nonlinear Sci. Series A *16*, 441–446.
97. Rosso, M., Moens, B., Leman, M., and Moumdjian, L. (2023). Neural entrainment underpins sensorimotor synchronization to dynamic rhythmic stimuli. Neuroimage *277*, 120226.
98. Barr, D.J., Levy, R., Scheepers, C., and Tily, H.J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. J. Mem. Lang. *68*, 255–278.
99. Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). Fitting Linear Mixed-Effects Models using lme4. Preprint at arXiv. https://doi.org/10.48550/arXiv.1406.5823.
100. Kay, M.M., Kay, M., and Wobbrock, J.O. (2020). Package "ARTool" (CRAN Repository). 2016. https://doi.org/10.1145/1978942.1978963.

CellPress
OPEN ACCESS

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Datasets	This paper	https://data.mendeley.com/datasets/24njbrmyjj/1
Software and algorithms		
Matlab 2019a	Mathworks	RRID:SCR_001622
RStudio	Comprehensive R Archive Network	RRID:SCR_000432

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Mattia Rosso (mattia.rosso@ugent.be).

Materials availability

This study did not generate new materials.

Data and code availability

Data and code have been deposited at Mendeley and are publicly available as of the date of publication Mendeley Data: https://doi.org/10.17632/24njbrmyjj.1.

Any additional information required to reanalyse the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Forty (N = 40) right-handed human participants took part in the study (28 females, 12 males; mean age = 31.42 years, standard deviation = 7.49 years). To ensure diversity in ancestry, race, and ethnicity, participants were recruited from a diverse and international pool. In order to control for gender bias in the interaction, they were divided into two gender-matched groups and randomly paired in twenty (N = 20) dyads. One dyad was excluded from dyadic analyses due to failure to comply with the instructions. Due to a technical problem in the video streaming during the procedure, one participant was excluded from the analysis of self-reported ownership. None of the participants had history of neurological, major medical or psychiatric disorders. All of them declared they were not professional musicians upon recruitment, although some of them had musical experience. None of the participants declared to know the assigned partner before the experiment. The study was approved by the Ethics Committee of Ghent University (Faculty of Arts and Philosophy) and informed written consent was obtained from each participant, who received a 20€ coupon as compensation for their participation.

METHOD DETAILS

Partners were sitting across the same table, facing one other. In the preparation phase, they were assisted in wearing a black cloth over the whole body and a long red glove over the right hand, with the purpose of rendering the visual scene as neutral as possible and minimizing individual differences related to personal clothing and skin texture. Before proceeding further, each individual participant underwent a period of familiarization with the finger-tapping task. Specifically, an auditory metronome was presented via in-ear plugs and he/she was instructed to tap the right index finger on a circular pad placed on the table. The experimenter showed how tapping was supposed to be performed, so that both partners would adopt a common style during the task. Pink noise played in the background alongside the metronomes, with the volume adjusted so each participant could clearly hear the metronome but not the feedback from their own tapping on the pad.

Participants were then equipped with HTC Vive Pro 2 headsets for immersive virtual reality (VR) environments, and underwent the standard calibration procedure as implemented by the manufacturer. Each set was connected to a different computer, running a Unity executable which took video input from a Logitech Brio Ultra HD Pro Business webcam (USB 3.0) and streamed it to the head-mounted display. The setup allowed to present an immersive photorealistic view of the right hand up to the forearm. The hand could be seen in either 1st or 2nd person perspectives, and could either pertain to one's own or to the partner, depending on the experimental condition. Before each condition had taken place, the factor Perspective was manipulated by placing two cameras above the shoulder of the participant (1P) or in front of the partner's hand (2P). The factor Ownership was manipulated by simply swapping the USB connection of the cameras to the respective computers, so that participants would perceive their own (Self) or the partner's (Other) hand. The resulting visual scenes can be seen in the details of Figure 1 from all levels of Perspective and Ownership across experimental conditions. Extensive testing prior to the beginning of the study resulted in an average video latency of 96ms for streaming 1080p video at a frame rate of 60Hz. Previous pilots and qualitative interviews with the participants revealed that the delay was barely perceivable, and in most cases not noticed at all.

In order to collect subjective reports of the sense of ownership over the visually perceived hand, we carried out the following procedure before starting each experimental condition. Both participants were asked to stay relaxed while the experimenter placed their right hand on a cardboard surface on the table, above the assigned tapping pad. They were instructed to watch for one minute the hand that would appear in the head-mounted display shortly thereafter, and to not move their own hand despite whatever would happen in the visual scene. From the moment the experimenter launched the video streaming, participants saw the hand lying still on a cardboard in front of them for 30 seconds (which hand and from which perspective depended on the experimental condition, as illustrated in the details of Figure 1). For the following 30 seconds, the experimenter applied synchronous touches on the back of the hand of both participants. Next, the visual scene went blank, and the participants were verbally asked the following question: "On a scale from 1 to 5, how much did you feel like the hand that you were seeing belonged to you?". The response was given in silence by raising the fingers of the left hand, in order not to bias the partner or induce motor activity in the stimulated hand.

What follows is the description of the 'drifting metronomes' paradigm, as originally described in Rosso et al.[6] Each partner was assigned to one pad and instructed to tap on it with the right index finger, synchronizing with an auditory metronome. The two metronomes slightly differed in tempo (1.67Hz and 1.64Hz), whereas the timbre remained the same. With the start of the two metronomes' tracks being aligned, the relative phase between metronomes started at 0° and steadily increased in regular steps of 5.6°. A full cycle took 39.008 seconds to be completed (65 and 64 clicks of the faster and slower metronome, respectively). Ten consecutive cycles were performed in each experimental condition. In conditions of informational coupling, participants were instructed to ignore their partner and to tap along with the assigned metronome. Participants' chairs were provided with an armrest, in order to exclude any tactile or proprioceptive coupling due to vibrations of the table resonating with finger taps.

A M-Audio® M-Track 8 soundcard was used to route independent audio channels to each participant via in-ear plugs. The average audio latency from tapping pad to earplug was 17 ms, with a standard deviation of 2 ms. Ableton Live 10 ® was used as the main interface for stimuli presentation, with 2 separate MIDI tracks triggering the metronome's audio sample. A Teensy 3.2 microcontroller was used as a serial/MIDI hub in the setup: tapping onsets were detected with 1ms resolution using analog input of strain gauge sensors installed inside the pads, while metronomes onsets were logged using MIDI messages originating from Ableton. Each class of events (metronomes 1 and 2, finger-taps of participants 1 and 2) was retrieved by means of a predefined ID number. Simultaneous EEG recordings were performed from both partners of the dyads during the whole experiment, but such data are not presented in the present paper. Additional data were collected prior and during the experiment. Prior to the experiment, demographic data were collected; the *Edinburgh inventory*[89] was administered to assess the right handedness of the participants; the 28-items version of the *Interpersonal Reactivity Index* (IRI)[61] was administered as a self-report of empathy and its subscales. During the breaks between experimental conditions, all participants provided subjective self-reports on different aspects of the task by expressing agreement on a scale from 1 ("Completely disagree") to 7 ("Completely agree") with a custom-made battery of 11 Likert items. Among these, the sense of ownership and sense of agency experienced during the task were measured by asking to rate the respective following items: "I felt like the hand that I was seeing belonged to me", and "I felt like it was me moving the hand that I was seeing".

Participants were monitored by the experimenters from behind curtains, where the visual scene of their headsets was visible on two separate screens. Dyad 7 was excluded from the analysis, given that one participant was unable to comply with instructions during Condition 4.

QUANTIFICATION AND STATISTICAL ANALYSIS

Pre-processing

Over the course of 10 consecutive metronomes' cycles, 650 and 640 tapping onsets were expected from the partners forming each dyad, for the total duration of 390 seconds. Onsets occurring < 350ms from the previous one were considered false positives and removed, since participants could occasionally push the pad for too long or accidentally lay their hand on it. Out of the whole sample, 59 false positives were removed, corresponding to 0.06% of all data points. The cleaned time series were then interpolated with a sine function at 1kHz sampling rate, providing an estimate of the oscillators' positions on its cycle with a temporal resolution of 1ms. The tap preceding the first metronome onset and the last tap following the last metronome onset were included in the interpolation. Afterwards, data points outside the boundaries of the metronomes time series were removed. Operationally, the procedure guaranteed equally sized time series without loss of data, which was a requirement for the application of joint recurrence quantification analysis (JRQA; see next paragraph). The modelling of systems of coupled oscillators in the context of joint finger-tapping studies conceptually supports the choice of interpolation.[6,65,83,90] Finally, time series were down-sampled by a factor of 4 to make computation of recurrence plots (RPs) computationally feasible. As shown in our previous report, results of JRQA are robust to the choice of the down-sampling factor.

Phase-space reconstruction

In accordance with Takens' embedding theorem,[91] we reconstructed the phase space of individual finger-tapping behaviors. This was done based on time-delayed copies of the input time series u_k, applying an embedding dimension m and a time delay τ.

$$\vec{x}(t) = \vec{x}_i = \left(u_i, u_{i+\tau}, \ldots, u_{i+(m-1)\tau}\right), \quad t = i\Delta t$$

where $\overrightarrow{x}(t)$ is the vector of reconstructed states in the phase-space at the time t. Optimal parameters for the time-delayed embedding were computed for each participant, for the time course of each single metronome's cycle in all experimental conditions. The resulting mean value was applied to all individual instances. The reason for this approach is that in order to compare the rate of recurrences across conditions at the group level, the embedding procedure must be consistent across participants (e.g., see,[92] for an example of parameter selections in a factorial design). We first selected the delay τ as the first local minimum of mutual information index[93] in function of delay. This approach minimized the time series self-similarity, extracting nearly orthogonal components and preventing the attractor from folding over itself.[94] Mean value of delay resulted to be $\tau = 7$. Next, we determined the number of embedding dimensions with the method of false nearest neighbor.[95] Specifically, we progressively unfolded the time series into higher dimensions until data points did not overlap spuriously, finding an optimal mean embedding of m = 2. Finally, in accordance with the literature, the maximum threshold for counting two neighboring points as recurrent was set at 10% of the maximal phase-space diameter.[58]

Joint recurrence plots (JRPs)

A recurrence plot $R_{i,j}$ is a square array used to represent and quantify recurrences of states in the phase space of a system.[96] For every point of the phase space trajectory:

$$\overrightarrow{x}_i(i = 1, \ldots, N; N = n - (m-1)\tau)$$

We tested whether it was close to another point of the trajectory $\overrightarrow{x}_j$ based on a neighborhood threshold. Individual recurrence plots were computed as follows:

$$R_{i,j}(\varepsilon) = \Theta(\varepsilon - \|x_i - x_j\|)$$

where ε is the neighborhood threshold, $\| \cdot \|$ is the Euclidean norm, representing the distance between two vectors, and Θ is the Heaviside step function. A square matrix was returned from each phase-space, containing 1s for all the instances where the distance $\| \cdot \|$ was smaller than the threshold ε, and 0s for remaining elements. A joint recurrence plot (JRP) was computed for each dyad by pair-wise overlapping partners' individual RPs, and keeping 1s only the instances where both plots contain a recurrence. Each JRP is in fact the Hadamard product of the recurrence plot of the first participant and the recurrence plot of the second participant. Computation of JRPs was carried out using the *crp toolbox* for Matlab ®.[58]

The 10 trials (i.e., the metronomes' cycles) of each experimental condition were aggregated by summing the respective JRPs of each trial. This resulted in a 2-D matrix for which every entry contained the amount of recurrences occurring in the corresponding region of the cycle, across all trials. Finally, a 1-D vector recurrence scores was obtained by looping over the columns of the matrix and summing the counts contained in the rows. This vector represents a density measure of the instances of coupled behavior over the course of the metronomes' cycle. The scale of these recurrence scores depend on the size of the JRPs and in turn on the embedding procedure, which makes it necessary to set the same parameters on the whole sample. In order to improve signal-to-noise ratio and avoid over-sampling in view of our statistical model, the resulting time series were reduced to 64 bins by averaging the recurrence score for equally sized, consecutive time periods. For this segmentation, interval size was equal to the slower metronome's increments, as they provided a regular subdivision intrinsic to the experimental trials. All processing steps presented were carried out in Matlab ®. Our approach was preferred over the version for JRQA based on moving windows, for the latter would act as a low-pass filter on our time series and hinder the interpretation of our results. Specifically, a moving window results in a phase distortion of the time series dependent on window size, and is thus not reliable in detecting attractor points over the attractor landscape. Since the procedure hereby described reproduced exactly the steps in Rosso et al.,[6] the content of the present paragraph is taken from the original work with the consent of the authors. Values for the embedding dimension and delay do differ from the original work, because they were optimized for the present dataset.

Statistical models

The recurrence score was used as response variable in a mixed-effects model with Coupling and Perspective as factors, and Time as a continuous predictor expressed with the indexes of the metronome's steps (from 1 to 64). Given the non-linear time-course observed in coupled conditions, we adopted the method of orthogonal polynomials including linear and quadratic functions of Time into our model.[6,59,97] Dyads and interactions between Dyads and the factors were modelled as random effects on all polynomial terms, to account for the individual variability in synchronization skills and individual susceptibility to coupling across the experimental manipulations. The random effects structure was used in order to minimize false alarm rates without substantial loss of power.[98] Informed by our previous study and by the inspection of empirical curves from the present dataset, we limited the polynomial model to the 2nd order as the most parsimonious solution. In this analysis framework, the intercept is considered a 'zero-order' polynomial, as it exhibits zero changes in any direction. Significant changes of direction indicate modulation by the temporal structure of the task. This allowed us to quantify the influence of attractor points, as the dyad deviated from the horizontal trajectory transitioning over expected critical regions.

The formula of the full model is the following:

$$Recurrence \sim (Time + Time^2) * Coupling * Perspective + (Time + Time^2 | Dyad) + (Time + Time^2 | Dyad : Coupling : Perspective)$$

Aligned rank transform (ART) ANOVA[60] was used to test the 2-way interaction between factors on the ratings of sense of ownership over the visually perceived hand, which as an ordinal response variable does not conform to the assumptions of a parametric factorial ANOVA. The same model was fit to the ratings of sense of ownership and sense of agency referring to the joint finger tapping task.

The formulas of the ART ANOVA models are the following:

$$Ownership \sim Coupling * Perspective$$

$$Ownership\_task \sim Coupling * Perspective$$

$$Agency\_task \sim Coupling * Perspective$$

The 3-way interactions of the IRI empathy subscales[61] with the factors were tested by fitting separate mixed-effects linear models for every subscale on the synchronization consistency (with the assigned metronome) and on the ratings of ownership. Subjects were modelled as random effects. Observations whose residuals deviated by more than 2 standard deviations from the mean where identified as outliers and removed from the analysis.

The formulas of the two linear models are the following:

$$R \sim Coupling * Perspective * IRI_{subscale} + (1|Subject)$$

$$Ownership \sim Coupling * Perspective * IRI_{subscale} + (1|Subject)$$

Statistical analyses were carried out in R (version 4.0.3). *lme4*[99] and ARTool[100] packages were used for model fitting.