Dyadic entrainment. An overwhelming and modality-dependent phenomenon.

Mattia Rosso*, Pieter J. Maes¹, Marc Leman¹

¹ IPEM Institute for Systematic Musicology. Ghent University - Department of art, music and theatre sciences.

Email (corresponding author): mattia.rosso@ugent.be
Abstract

Rhythmic joint coordination is ubiquitous in daily-life human activities. In order to coordinate their actions towards shared goals, individuals need to co-regulate their timing and move together at the collective level of behavior. Remarkably, basic forms of coordinated behavior tend to emerge spontaneously as long as two individuals are exposed to each other’s rhythmic movements.

The present study investigated the dynamics of spontaneous dyadic entrainment, and more specifically how they depend on the sensory modalities mediating informational coupling. By means of a novel interactive paradigm, we showed that dyadic entrainment systematically takes place during a minimalistic rhythmic task despite explicit instructions to ignore the partner. Crucially, the interaction was organized by clear dynamics in a modality-dependent fashion. Our results showed highly consistent coordination patterns in visually-mediated entrainment, whereas we observed more chaotic and more variable profiles in the auditorily-mediated counterpart.

The proposed experimental paradigm yields empirical evidence for the overwhelming tendency of dyads to behave as coupled rhythmic units. In the context of our experimental design, it showed that coordination dynamics differ according to availability and nature of perceptual information. Interventions aimed at rehabilitating, teaching or training sensorimotor functions can be ultimately informed and optimized by such fundamental knowledge.

Introduction

Humans are in remarkable ways influenced by what they hear and see, especially when they perceive other people moving. The exposure to rhythmic behaviors attracts individuals to fall in sync with one another, and the phenomenon can be observed at the level of dyads, small-groups and large crowds. For instance, when two people walk, run or dance together, they show a natural tendency to synchronize their steps. When an audience applauds at the end of the concert, the auditory scene goes through intermittent periods of synchronized clapping. When marching bands cross each other on the street, it requires conscious effort for them to avoid getting rhythmically entrained.

Rhythmic joint coordination is ubiquitous in daily-life human activities, from the simplest form of unintentional synchronization to complex tasks such as musical performance. The way different rhythms organize into orderly coordinated patterns over time implies co-regulated actions in response to rhythms produced by other agents. Characteristic for this co-regulation is that action patterns are often spontaneously attracted towards stable synchronization, via a process called entrainment. ‘Coordination dynamics’ has been proposed as a theoretical framework to understand the generic organizational principles that underlie the coordination of coupled rhythmic units. Central to the theory is that the coordination of multiple units over time is the result of a dynamic balance of keeping the own rhythm (competition) and moving together on the collective level (cooperation). Such idea resonates in earlier work by von Holst, who termed the ‘maintenance effect’ versus the ‘magnet effect’ in his study of coordination in animal behavior.
When intrinsic and external rhythms are incongruent, rhythmic units engage in a dynamic competition-cooperation process. When it comes to rhythmic interactions within a dyad, this dynamic balance can be influenced by manifold variables, such as preferred tempo \(^{12}\), empathy \(^{13}\) and social factors \(^{14}\). However, among all possible sources of variability, informational coupling between individuals stands out as the most fundamental condition for interpersonal coordination. The essential minimal requirement for human interactions to occur is that individuals mutually exchange information through one or more sensory channels. Although experimental interactive scenarios necessarily imply a determined sensory modality \(^{15-17}\) and possibly cross-modal interactions \(^{18}\), the role of competing sensory modalities remains under-looked in the literature. To our knowledge, whether and how coordination dynamics depend on the nature of available perceptual information is yet to be investigated.

The aim of the present study is to develop a dynamic experimental paradigm, and apply it for studying modality dependencies of competition-cooperation dynamics in dyadic entrainment. The paradigm involves a unimanual tapping task where two individuals are exposed to minimally different tempi and reciprocally coupled in visual and auditory modalities across experimental conditions. The procedure implies a tension between competition and cooperation processes in dyadic interaction, and is meant to return an empirical layout of its dynamics across sensory modalities.

The core idea underlying the proposed experimental paradigm is to induce a competition between the intrinsic rhythm of an individual’s actions and the rhythm produced by another person. During a unimanual tapping task, two partners are exposed to two metronomes in different modalities (visual and auditory). The metronomes assigned to the partners slightly differ in frequency, resulting in a gradual increasing (0 to \(\pi\) rad) and decreasing (\(\pi\) to 0 rad) of absolute relative phase over multiple cycles. Accordingly, for each person, a corresponding cyclical de-phasing of the assigned metronome and partner’s metronome unfolds over time. With this simple expedient, we implemented a completely predictable system of two oscillators which deterministically revisits the same states throughout 10 consecutive cycles. An audio file showing metronomes’ behavior in available in the Supplementary materials.

When the partners are informationally coupled, they are exposed to two competing rhythms. The first one is the intended rhythm: participants are explicitly instructed to synchronize to their assigned metronome. The second rhythm is the coupling rhythm, which is produced by the partner and perceived in another sensory modality. Thus, when the members of the dyad follow their own auditory metronome, they are exposed to visual information from the partner (visual coupling). Alternatively, when the members of the dyad follow their own visual metronome, they are exposed to auditory information from the partner (auditory coupling). The same task performed in absence of informational coupling provides the control conditions for each modality, and a baseline to assess the impact of coupling on spontaneous dyadic entrainment. In presence of informational coupling, a cross-modal incongruency takes place between intended and coupling rhythms. The experimental design is schematized and described in Figure 1.

In the framework of coordination dynamics, the concept of relative phase \((\phi)\) functions as a collective variable to describe co-regulation of timing. When competition dominates the interaction, rhythmic units may fully keep their own intrinsic frequency, while they may be fully attracted towards each other as cooperation prevails. The former case would typically result in an evenly spread distribution across \(\phi\) (0 to 2\(\pi\) radians). In the latter case, the distribution would exhibit peaks over the so-called attractor regions, where rhythmic units are locked in temporally stable state. In human behavior, the intrinsic dynamics of the collective variable \(\phi\) are well characterized by a layout of attractive and repelling fixed points, wherein in-phase (0 radians) and anti-phase (\(\pi\) radians) are the dominant stable modes of coordination. The Haken-Kelso-Bunz (HKB) model mathematically describes how phase transitions from one spatiotemporal pattern to another take place when the system is pushed beyond its equilibrium state \(^{19}\). The model points at the in-phase mode as the point where the system ultimately tends to reach stability.

Crucially, the behavior of the ‘drifting metronomes’ recursively guided the dyads through the whole range of \(\phi\) values – and over its respective attractor regions. Based on a joint recurrence analysis
of the system’s dynamics, we computed a recurrence score as a metric of dyadic coupling strength and tracked its evolution over time (for a visual representation of the processing pipeline, see Figure 2 & 3). By these means, we obtained an empirical attractor layout descriptive of the competition-cooperation dynamics within the dyads. The cooperation process (i.e., entrainment to the partner’s coupling rhythm) was expected to dominate the conflict around in-phase (0 radians) and anti-phase (π radians) points, leading to significantly higher recurrence score in presence of informational coupling and a significant modulation around attractor regions.

Given the time-varying nature of the recurrence score as a response variable, we needed a method to model its local variations over the course of the metronomes’ cycles. We opted then for a statistical model which allowed us to assess the effects of our experimental manipulations over the entire unfolding of the interaction, and to avoid segmentation and multiple comparison while respecting the temporal structure of the data. In addition to the inferences made possible by linear models based on intercepts, such as ANOVAs, our solution allowed to assess the effect of categorical predictors on the temporal profile of the response variable.

**Results**

*Growth curve analysis* was used to analyze the evolution of recurrence score over the course of the expected relative phase between partners (i.e., from 0 to 2π). The curves were calculated as the average of the 10 consecutive metronomes cycles in each experimental condition, and modeled with 2nd order orthogonal polynomials and fixed effects of Coupling and Modality on the polynomial terms. The uncoupled (control) conditions were treated as the baseline, and parameters were estimated for the coupled conditions. Similarly, the audio conditions were treated as a baseline, and parameters were estimated for the visual conditions.

In the context of orthogonal polynomials, a flat line can be considered a pure intercept and a ‘zero-order’ polynomial, in the sense that it exhibits zero changes in any direction. If the recurrence score was time-invariant, it would appear as a flat line indicating complete dominance of the competition process between rhythmic units, which would pursue their individual intended rhythms. On the other hand, significant changes of direction would indicate that the interaction is systematically modulated by the temporal structure of the task: around attractor points, the cooperation process would take over and the rhythmic units would be attracted towards the coupling rhythm. Figure 4 shows how orthogonal polynomials are suited for modelling local variations of the response variable around attractor points.

Orthogonal polynomials in the model were limited to the 2nd order: based on our hypotheses, we expected quadratic and quartic terms to capture the effects of in-phase and anti-phase attractor points. However, following data inspection and analysis of residuals, we concluded that the quadratic term alone was enough to capture the relevant effect of our experimental manipulation. Figure 5 shows the data with descriptive statistics across experimental conditions, and a comparison between 2nd order and 4th order models: arguably, the better fit of the quartic term comes with a major cost in terms of model parsimony and interpretation.

Our model also included random effects of Dyads on all polynomial terms, and their interactions with the factors: the random effects structure was maximized in order to minimize false alarm rates without substantial loss of power. As expected, we found a significant main effect of Coupling (*Estimate* = 42.809, *SE* = 18.142, *p* = 0.018) indicating an overall increase of the recurrence score in presence of informational coupling as compared to the uncoupled control conditions, independently from the sensory modality. We also report a significant interaction effect of Coupling and Modality on the orthogonal polynomial terms of Time (Linear: *Estimate* = -84.579, *SE* = 30.867, *p* = 0.006; Quadratic: *Estimate* = 190.320, *SE* = 36.973, *p* < 0.001), meaning that the parameter estimates in coupled conditions significantly differ across visual and auditory modalities. Given the observed major variability in the Auditory level and following inspection of the residuals, we fitted the same model on the natural log-transformed response variable for it provided a considerably better fit, and still found significant main effect of Coupling (*Estimate* = 0.237, *SE* = 0.117, *p* = 0.042) and interaction effects on the
polynomial terms of Time (Linear: Estimate = -0.534, SE = 0.222, p = 0.016; Quadratic: Estimate = 1.257, SE = 0.299, p < 0.001). Table 1 shows the fixed effects parameter estimates and their standard errors for log-transformed data, along with p-values estimated using the normal approximation for the t-values.

| Predictors        | Estimates | SE   | p     |
|-------------------|-----------|------|-------|
| Time              | -0.0375   | 0.113| 0.739 |
| Time²             | -0.360 *  | 0.150| **0.017** |
| Modality          | 0.155     | 0.117| 0.184 |
| Coupling          | 0.237 *   | 0.117| **0.042** |
| Time:Modality     | -0.021    | 0.157| 0.893 |
| Time²:Modality    | 0.269     | 0.212| 0.202 |
| Time:Coupling     | 0.011     | 0.157| 0.942 |
| Time²:Coupling    | 0.055     | 0.212| 0.793 |
| Modality:Coupling | 0.118     | 0.165| 0.475 |
| Time:Modality:Coupling | -0.534 * | 0.222| **0.016** |
| Time²:Modality:Coupling | 1.257 *** | 0.299| < **0.001** |

Table 1. Model summary.

In the context of growth curve analysis, the parameter estimates provide a measure of effect size of straightforward interpretation for linear and non-linear changes overtime as long as the order not too high: above the 4th order, results become hardly interpretable and the risk of over-fitting the data increases. With the interaction effect of Coupling and Modality factors on the polynomial terms, we could quantify modality-specific effects of informational coupling on the evolution of the recurrence score. The results strongly support our predictions of spontaneous dyadic entrainment in presence of informational coupling and, crucially, show that its temporal dynamics are modality-dependent.

Discussion

The present study investigated the dynamics of dyadic entrainment, and more specifically how they depend on the sensory modality mediating informational coupling. An experimental paradigm was designed to dynamically manipulate the expected relative-phase between participants, while informational coupling was manipulated in visual and auditory modalities across conditions. At the global level, we wanted to prove that when intended and coupling rhythms are at odds, a
spontaneous co-regulation of timing occurs within the dyad at the expenses of the individual intended rhythms. At the local level, we wanted to infer the configuration of a dyadic basin of attraction, by modelling the dyadic cooperation process around attractor points. Ultimately, it was possible to evidence a crucial modality-dependency of dyadic entrainment.

In the first place, we reported that the mere presence of informational coupling led to spontaneous entrainment in dyads engaged in a rhythmic task. As long as two partners perceived each other, their rhythms were 'attracted' towards each other at the expenses of the synchronization with the respectively assigned metronomes. The effect was quantified as a significant increase of the average recurrence score, compared to the baseline provided by the de-coupled conditions. Remarkably, such manifestation of self-organizing behavior emerged in the context of a task where subjects were explicitly instructed to ignore each other's actions and to pursue independent uncoupled behaviors cued by their own reference. Even though the observation of spontaneous dyadic entrainment was reported in self-paced, cued and joint improvisation tasks, our procedure explicitly implemented rhythmic incongruency to investigate the competition-cooperation dynamic balance latent to the system's dynamics. When competition prevails, rhythmic units manage to keep the assigned (intended) rhythm, whereas when cooperation takes over, rhythmic units undergo a (spontaneous) co-regulation of timing and move together on a collective level of coordinated behavior. Dyadic entrainment occurred regardless of the sensory modality mediating the informational coupling.

Beyond a global measure of coupling, the continuous de-phasing implemented in the 'drifting metronomes' allowed us to quantify local variations of recurrence score as a function of the expected relative phase between partners. The balance between cooperation and competition processes is dynamic, in the sense that it varies over time and is constrained by the intrinsic dynamics of the system. Dyadic entrainment was expected to occur more strongly and more consistently around attractor regions, located around the in-phase and anti-phase points of the metronomes' cycles. With our approach, we achieved to return an empirical 'attractor landscape' and to contrast it across conditions of visual and auditory coupling. Our findings strongly point towards a modality-dependency of dyadic entrainment (see Figure 5). The interaction effect of Coupling and Modality on the linear and quadratic terms of Time statistically confirms that the temporal modulation is an exclusive property of visually-mediated coupling, with a remarkable consistency across dyads. In such condition, the recurrence score exhibits maximum values around the in-phase point at the start of the metronomes cycle, gradually decreases as the dyad leaves the attractor zone, reaches its minimum passed the anti-phase point, and finally increases towards maximum values as it enters the in-phase zone at the end of the cycle. We report here an effect of directionality in the visually coupled condition: dyads slowly de-couple as they leave the in-phase point, whereas they couple more rapidly as they approach the same point. The negative slope of the linear term indicates that dyads tend to be more coupled in the first portion of the cycle, namely before passing the anti-phase point. To the best of our knowledge, such phenomenon has not been reported in the literature before.

Differently, in condition of auditory coupling the recurrence score did not exhibit any significant variation over time: the measure was locally independent from the proximity of attractor points. Nevertheless, we still observed a significant global increase of recurrences compared to the baseline. Previous evidence shows that phase shifts in auditory sequences elicit a phase-correction response in participants instructed to tap in synchrony to isochronous sequence of flashing lights. Auditory-driven phase corrections occur even when shifts are highly irregular and at different cross-modal relative phases, with overall high interindividual variability. In the context of our paradigm, sounds generated by the partner exerted the same attraction among individual rhythms regardless of the expected relative-phase.

Finally, both control conditions resulted in a flat horizontal line below the average of the respective coupled conditions. It is worth noticing the higher variability in Audio conditions, arguably due to the task of synchronizing to visual metronomes: from lower recurrence score in the control condition we can conclude that dyads performed globally worse in synchronizing to visual cues as compared to auditory cues, which is in line with previous evidence on modality-dependent synchronization skills, and with participants' self-reports about the perceived difficulty of the
task. Taken together, our findings support our hypothesis of modality-dependent dynamics, suggesting that auditorily-mediated entrainment is considerably less sensitive to the basin of attraction when compared to its visual counterpart.

Although interpersonal coordination can be described at the collective level as a self-organizing process, 9,25 humans ultimately entrain to each other via mutual adaptation of individual action-perception loops 16,17. Therefore, an exhaustive account for dyadic entrainment should be capable of bridging the two levels, considering how coordination dynamics depend on the processing of information available in the interaction. Predictive coding 36–40 was first formulated as a general theory for neural computation, and more recently put forward as the potential mechanism underpinning coordination dynamics in dyadic behavior 41. The theory states that the brain is constantly engaged in an optimization process based on Bayesian inference: information sampled from the environment is compared to prior evidence, and the optimization consists of minimizing prediction errors. In the context of interpersonal coordination, such a putative mechanism could underpin the overt manifestations of dyadic entrainment by engaging the action systems of the rhythmic units involved. Let us consider the prediction error as the phase difference between executed and observed movements: as both units minimize the mismatch between the representations of observed and own motor behavior, the dyadic system tends towards a collective minimization of free energy in a stable attractor state 41. Such interpretation of the dyadic attractor is supported by our results in the context of visually-mediated coupling: as the relative phase between partners increases, the cooperation process between rhythmic units loses strength in favor of competition. In other terms, the tendency of the partners to lock in a stable coordinative state depends on the energy required to correct for the prediction error. As the error increases, the correction becomes more effortful and it becomes easier for the units to pursue independent behaviors.

Crucially, such explanation does not hold for auditorily-mediated coupling. As distinct sensory systems have their unique interface with motor and timing systems 37,42,43, it is necessary to interpret the results in light of individual information processing. There is arguably a crucial difference in the nature of the perceptual information available to the partners across different modalities. When the coupling is visually mediated, a stream of kinematic information is continuously sampled from the partner's actions 44, such that the predictive models can be constantly updated at every stage of movement execution. On the other hand, in case of auditory coupling the prediction of the partner's tap is solely based on temporal information about the previous taps, since partners do not have online access to the reciprocal kinematic information. In such condition, the cooperation process still dominates the interaction but no consistent pattern emerges in terms of dynamics as quantified by our method. We propose that the availability of kinematic information through informational coupling is a crucial discriminant that might explain the observed modality-dependence of dyadic entrainment.

A final question arises spontaneous: how would the dyad behave if kinematic information was embedded in a sound envelope, and conveyed to the partners via the auditory channel? One limitation of the current work is that we cannot conclude whether dyadic entrainment is modality-dependent in the strict sense or rather dependent on availability of kinematic information. Given the central role attributed to kinematics in predictive accounts for action perception 44, it is of particular interest to investigate whether continuous sonification of movement parameters 45 would support the same dynamics observed in presence of visual coupling. Previous work shows that presenting visual stimuli in a way that indicates movement over time, e.g., apparent hand motion 46,47 or a bouncing ball 36,42,48, behavioral outcomes of predictive entrainment are improved by the continuous stream of information. We hypothesize that a movement sonification strategy based on this principle could translate the kinematic advantage into auditorily-mediated dyadic interactions, resembling the dynamics here described in condition of visual coupling.

The findings and the questions raised from the present study might have relevant implications for optimization of interventions aimed at rehabilitating, teaching or training sensorimotor functions. Fundamental knowledge about dyadic interactions is what will ultimately inform such optimization. The evidence from the present study points at the advantages of the physical presence of a teacher.
(or a therapist, or a trainer) for guiding rhythmic movements, possibly integrating kinematic information via sonification strategies. In conclusion, we want to highlight that the ‘drifting metronomes’ procedure is meant as a methodological contribution for the investigation of interpersonal coordination, in the hope that the community can build upon it with new research questions and experimental designs. We point at its adoption with simultaneous dual-electroencephalography recordings, computational simulations and replication on pathological populations as potential sources of insight into the fundamentals of dyadic interactions.

Methods

Participants. Twenty-eight right handed participants took part in the study (18 females, 10 males; mean age = 29.07 years, standard deviation = 5.73 years). They were randomly paired in fourteen (N = 14) gender-matched dyads, in order to control for any gender bias in the interaction. None of them had history of neurological, major medical or psychiatric disorders. All of them declared not to be professional musicians upon recruitment, although some of them had musical experience. With the exception of one dyad, all participants declared not to know the assigned partner from before the experiment. The experiment 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 15€ coupon as economic compensation for their participation.

Experimental apparatus and procedure. The two partners sitting across the same table, facing each other. Chairs were provided with an armrest in order to exclude any tactile or proprioceptive coupling due to the table vibrating. Each partner was assigned to one pad and instructed to tap on it with the right index finger synchronizing with a metronome. As represented and described in Figure 1, the metronome was either an auditory cue or a flickering light, depending on the experimental condition. Each partner was presented with a different metronome tempo (1.67Hz and 1.64Hz). Aligning the start of the two tracks, the relative phase between the metronomes started at 0º and steadily increased in regular steps of 5.6º. A full cycle took 39 seconds to be completed. 10 consecutive cycles were performed in each experimental condition. In conditions of informational coupling, participants were instructed to ignore the partner and tap along with the assigned metronome. A M-Audio® M-Track 8 soundcard was used to route independent audio channels to each participant via in-ear plugs. Ableton Live 10® was adopted as main interface for stimuli presentation, to sonify in real-time the participants’ finger-taps and to route them. The same MIDI tracks were used to control the metronomes across conditions, by either triggering an audio sample or a flickering LED. Volume was adapted to every participant before the start of the experiment, and pink noise was regulated up to the point of suppressing any sound but the auditory stimulation presented via Ableton. Participants were monitored on-line by means of a USB-camera, to make sure they complied with the instructions. No dyad was excluded from the analysis.

A Teensy 3.2 microcontroller was used as serial/MIDI hub in the setup. It was used to detect tapping onsets with 1 ms resolution, based on the analog input from the piezo sensors installed inside the pads. Every time a metronome onset was presented to a participant, a MIDI message was sent to the Teensy device to log the metronomes timeseries and to control the voltage of the LEDs when needed.

Simultaneous EEG recordings were performed from both partners of the dyads during the whole experiment, but such data were not presented in the present paper. Additional data were collected prior and during the experiment. Prior to the experiment, demographical data were collected; the Edinburgh inventory was administered to assess the right handedness of the participants; the Interpersonal Reactivity Index (IRI) was administered as self-report of empathy and its subscales; the preferred spontaneous tempo was calculated via 30 seconds of self-paced finger-tapping on a dedicated smartphone app (www.beatsperminuteonline.com). 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.

Data Analysis

Pre-processing. Our raw data consisted of timestamps logged from the Teensy controller with 1ms resolution for 10 consecutive metronome cycles (390 seconds in total), and an associated ID for each partner and each metronome. As only form of cleaning, we removed onsets occurring <350ms from the previous one: false positives could occasionally be recorded when a participant pushed the pad for too long or accidentally laid the hand on it. The cleaned timeseries 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. Conceptually, the choice of sinusoidal interpolation was supported by recent work on modelling of systems of coupled oscillators in joint finger-tapping studies. Operationally, it guaranteed that all timeseries match in size without any loss of data, which was a requirement for the next steps of our analysis. Timeseries were finally down-sampled by a factor of 40 to make the computation of RPs computationally feasible. Different orders of down-sampling were tested to make sure that the results do not depend on this choice.

Phase-space reconstruction. The optimal parameters for the time-delayed embedding were computed for each participant, for the time course of each single metronomes 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, for an example of parameter selections in a factorial design). We first selected the delay tau of the timeseries (τ) as the first local minimum mutual information index as a function of delay: this approach minimizes the timeseries self-similarity, extracting nearly orthogonal components and preventing the attractor from folding over itself. The mean value of τ resulted to be 7. Next, we determined the number of embedding dimensions m with the method of false nearest neighbor: we progressively unfolded the time series into higher dimensions until the data points did not overlap spuriously, finding a mean m = 3. Finally, in order to determine a maximum threshold for counting two neighboring points as recurrent, we selected a radius of 10% of the maximal phase-space diameter.

Joint Recurrence Plots (JRPs). Individual recurrence plots were computed as follows:

\[
R_{ij}(\varepsilon) = \Theta(\varepsilon - \|x_i - x_j\|)
\]

where \(\varepsilon\) is the neighborhood threshold, \(\| \cdot \|\) is the Euclidean norm, and \(\Theta\) is the Heaviside step function. A square matrix was returned from each shadow-manifold in the phase-space, containing 1s for all the instances where the distance \(\| \cdot \|\) was smaller than the threshold \(\varepsilon\), and 0s for the remaining elements. The distance was computed with the method of maximum norm. A joint recurrence plot (JRP) was computed for each dyad by overlapping the individual JRPs of the partners pair-wise, and keeping as 1 only the instances where both plots contain a recurrence. The computation of the JRPs was carried out with the crp toolbox for Matlab. The 10 trials (i.e., the metronome cycles) of each experimental condition were aggregated by summing the respective JRPs such that every cell of the 2-D matrix contains the count of times that a recurrence occurs in the same point of the cycle. Finally, we looped over the columns of the matrix summing all the counts contained in the rows, obtaining a 1-D vector recurrence scores which represents a density measure of the instances of coupled behavior over the course of the cycle. The scale of the recurrence score depends 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 stabilize our response variable and avoid over-sampling in view of our statistical model, the resulting timeseries were divided in 64 segments averaging the recurrence score. For the
division, we chose the intervals determined by the steps of the slower metronome, as they provide an intrinsic regular subdivision of the experimental runs. All the steps presented so far were carried out in Matlab® (see Figure 2, for a schematic representation). Our approach was preferred over the “windowed” version for JRQA, for the latter acted as a low-pass filter on our timeseries, making it impossible to interpret our results. The resulting phase-shift in the timeseries was dependent on the choice of the window size, and hence not reliable in detecting attractor points over the landscape.

**Statistical model.** The recurrence score was used as response variable in a mixed-effects model with Modality and Coupling 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 the ‘Visual Coupling’ condition, we adopted the method of orthogonal polynomials 23 including linear and quadratic functions of Time into our model (see Figure 5). Dyads and interactions between Dyads and Factors were modelled as random effects 24 to account for the individual variability in synchronization skills and individual susceptibility to coupling across the experimental manipulations. The formula of the full model up to the 2nd polynomial order is the following:

\[
\text{Recurrence} \sim (\text{Time} + \text{Time}^2) \ast \text{Modality} \ast \text{Coupling} + (\text{Time} + \text{Time}^2 \mid \text{Dyad}) + (\text{Time} + \text{Time}^2 \mid \text{Dyad: Modality: Coupling})
\]

Statistical analyses were carried out in R (version 4.0.3); model fitting was performed with the lme4 package 60. All methods were carried out in accordance with relevant guidelines and regulations, based on the references provided in the respective paragraphs.

**Acknowledgments**

The present study was funded by Bijzonder Onderzoeksfonds (BOF) from Ghent University (Belgium). The hardware of the recording device was built by Ivan Schepers. Figure 1 was illustrated by Kevin Smink. Professor Peter Vuust and Dr. Ole Adrian Heggli from the Center for Music in the Brain (Aarhus University, Denmark) provided precious advice and feedback for the experimental design.

**Author Contributions**

M.R. conceptualized, designed and implemented the study. P.J.M. supervised and contributed to study design and data processing. M.L. supervised and contributed to study design and statistical modelling. M.R. collected, processed and analyzed the data. All authors wrote the main manuscript text.

**Competing interests**

The authors declare no competing interests.

**Data availability**
All data generated during this study are included in this published article (and its Supplementary materials), together with the scripts used to analyze them.

References

1. Zhang, M., Kelso, J. A. S. & Tognoli, E. Critical diversity: Divided or united states of social coordination. *PLoS One* **13**, e0193843 (2018).

2. Néda, Z., Ravasz, E., Brechet, Y., Vicsek, T. & Barabási, A. L. The sound of many hands clapping. *Nature* **403**, 849–850 (2000).

3. Clayton, M., Dueck, B. & Leante, L. *Experience and Meaning in Music Performance*. (Oxford University Press, 2013).

4. Schmidt, R. C. & Richardson, M. J. Dynamics of Interpersonal Coordination. in *Coordination: Neural, Behavioral and Social Dynamics* (eds. Fuchs, A. & Jirsa, V. K.) 281–308 (Springer Berlin Heidelberg, 2008).

5. Keller, P. Ensemble performance: Interpersonal alignment of musical expression. in *Expressiveness in music performance: Empirical approaches across styles and cultures*, (pp (ed. Fabian, D.) vol. 383 260–282 (Oxford University Press, xxx, 2014).

6. Phillips-Silver, J., Aktipis, C. A. & Bryant, G. A. The ecology of entrainment: Foundations of coordinated rhythmic movement. *Music Percept.* **28**, 3–14 (2010).

7. Schöner, G. Learning and recall in a dynamic theory of coordination patterns. *Biol. Cybern.* **62**, 39–54 (1989).

8. de Guzman, G. C. & Tognoli, E. The human dynamic clamp as a paradigm for social interaction. *Proceedings of the* (2014).

9. Kelso, J. A. S. *Dynamic patterns: the self-organization of brain and behavior*. (MIT Press, 1995).

10. Scott Kelso, J. A., Engstrom, D. A. & Engstrom, D. *The Complementary Nature*. (MIT Press, 2006).

11. Von Holst, E. *The behavioural physiology of animals: the selected papers of Erich von Holst*. (Methuen, 1973).
12. Schmidt, R. C. & Turvey, M. T. Phase-entrainment dynamics of visually coupled rhythmic movements. *Biol. Cybern.* **70**, 369–376 (1994).

13. Novembre, G., Mitsopoulos, Z. & Keller, P. E. Empathic perspective taking promotes interpersonal coordination through music. *Sci. Rep.* **9**, 12255 (2019).

14. Tschacher, W., Rees, G. M. & Ramseyer, F. Nonverbal synchrony and affect in dyadic interactions. *Front. Psychol.* **5**, 1323 (2014).

15. Schmidt, R. C., Carello, C. & Turvey, M. T. Phase transitions and critical fluctuations in the visual coordination of rhythmic movements between people. *J. Exp. Psychol. Hum. Percept. Perform.* **16**, 227–247 (1990).

16. Konvalinka, I., Vuust, P., Roepstorff, A. & Frith, C. D. Follow you, follow me: continuous mutual prediction and adaptation in joint tapping. *Q. J. Exp. Psychol.* **63**, 2220–2230 (2010).

17. Heggli, O. A., Konvalinka, I., Kringelbach, M. L. & Vuust, P. Musical interaction is influenced by underlying predictive models and musical expertise. *Sci. Rep.* **9**, 11048 (2019).

18. Miyata, K., Varlet, M., Miura, A., Kudo, K. & Keller, P. E. Modulation of individual auditory-motor coordination dynamics through interpersonal visual coupling. *Sci. Rep.* **7**, 16220 (2017).

19. Haken, H., Kelso, J. A. & Bunz, H. A theoretical model of phase transitions in human hand movements. *Biol. Cybern.* **51**, 347–356 (1985).

20. Marwan, N., Carmen Romano, M., Thiel, M. & Kurths, J. Recurrence plots for the analysis of complex systems. *Phys. Rep.* **438**, 237–329 (2007).

21. Demos, A. P. & Chaffin, R. Removing obstacles to the analysis of movement in musical performance: Recurrence, mixed models, and surrogates. *The Routledge companion to embodied music interaction* 341–349 (2017).

22. Demos, A. P., Chaffin, R. & Logan, T. Musicians body sway embodies musical structure and expression: A recurrence-based approach. *Music Sci.* **22**, 244–263 (2018).

23. Mirman, D. *Growth Curve Analysis and Visualization Using R*. (CRC Press, 2017).

24. Barr, D. J., Levy, R., Scheepers, C. & Tily, H. J. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J. Mem. Lang.* **68**, 255–278 (2013).
25. Tognoli, E., Zhang, M., Fuchs, A., Beetle, C. & Kelso, J. A. S. Coordination Dynamics: A Foundation for Understanding Social Behavior. *Front. Hum. Neurosci.* **14**, 317 (2020).

26. Lorås, H., Aune, T. K., Ingvaldsen, R. & Pedersen, A. V. Interpersonal and intrapersonal entrainment of self-paced tapping rate. *PLoS One* **14**, e0220505 (2019).

27. Okano, M., Shinya, M. & Kudo, K. Paired Synchronous Rhythmic Finger Tapping without an External Timing Cue Shows Greater Speed Increases Relative to Those for Solo Tapping. *Sci. Rep.* **7**, 43987 (2017).

28. Kimura, K., Ogata, T. & Miyake, Y. Effects of a partner’s tap intervals on an individual’s timing control increase in slow-tempo dyad synchronisation using finger-tapping. *Sci. Rep.* **10**, 8237 (2020).

29. Noy, L., Dekel, E. & Alon, U. The mirror game as a paradigm for studying the dynamics of two people improvising motion together. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 20947–20952 (2011).

30. Noy, L., Levit-Binun, N. & Golland, Y. Being in the zone: physiological markers of togetherness in joint improvisation. *Front. Hum. Neurosci.* **9**, 187 (2015).

31. Varlet, M., Nozaradan, S., Nijhuis, P. & Keller, P. E. Neural tracking and integration of “self” and “other” in improvised interpersonal coordination. *Neuroimage* **206**, 116303 (2020).

32. Tuller, B. & Kelso, J. A. Environmentally-specified patterns of movement coordination in normal and split-brain subjects. *Exp. Brain Res.* **75**, 306–316 (1989).

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

34. Kato, M. & Konishi, Y. Auditory dominance in the error correction process: a synchronized tapping study. *Brain Res.* **1084**, 115–122 (2006).

35. Repp, B. H. & Penel, A. Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychol. Res.* **68**, 252–270 (2004).

36. Iversen, J. R., Patel, A. D., Nicodemus, B. & Emmorey, K. Synchronization to auditory and visual rhythms in hearing and deaf individuals. *Cognition* **134**, 232–244 (2015).
37. Comstock, D. C., Hove, M. J. & Balasubramaniam, R. Sensorimotor Synchronization With Auditory and Visual Modalities: Behavioral and Neural Differences. *Front. Comput. Neurosci.* **12**, 53 (2018).

38. Friston, K. A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **360**, 815–836 (2005).

39. Friston, K. The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* **11**, 127–138 (2010).

40. Friston, K., Mattout, J. & Kilner, J. Action understanding and active inference. *Biol. Cybern.* **104**, 137–160 (2011).

41. Koban, L., Ramamoorthy, A. & Konvalinka, I. Why do we fall into sync with others? Interpersonal synchronization and the brain’s optimization principle. *Soc. Neurosci.* **14**, 1–9 (2019).

42. Hove, M. J., Iversen, J. R., Zhang, A. & Repp, B. H. Synchronization with competing visual and auditory rhythms: bouncing ball meets metronome. *Psychol. Res.* **77**, 388–398 (2013).

43. Hove, M. J., Fairhurst, M. T., Kotz, S. A. & Keller, P. E. Synchronizing with auditory and visual rhythms: an fMRI assessment of modality differences and modality appropriateness. *Neuroimage* **67**, 313–321 (2013).

44. Kilner, J. M., Friston, K. J. & Frith, C. D. Predictive coding: an account of the mirror neuron system. *Cogn. Process.* **8**, 159–166 (2007).

45. Maes, P.-J., Buhmann, J. & Leman, M. 3Mo: A Model for Music-Based Biofeedback. *Front. Neurosci.* **10**, 548 (2016).

46. Hove, M. J. & Keller, P. E. Spatiotemporal relations and movement trajectories in visuomotor synchronization. *Music Percept.* **28**, 15–26 (2010).

47. Hove, M. J., Spivey, M. J. & Krumhansl, C. L. Compatibility of motion facilitates visuomotor synchronization. *J. Exp. Psychol. Hum. Percept. Perform.* **36**, 1525–1534 (2010).

48. Gan, L., Huang, Y., Zhou, L., Qian, C. & Wu, X. Synchronization to a bouncing ball with a realistic motion trajectory. *Sci. Rep.* **5**, 11974 (2015).
49. Liu, D. et al. Interactive Brain Activity: Review and Progress on EEG-Based Hyperscanning in Social Interactions. *Front. Psychol.* 9, 1862 (2018).

50. Heggli, O. A., Cabral, J., Konvalinka, I., Vuust, P. & Kringelbach, M. L. A Kuramoto model of self-other integration across interpersonal synchronization strategies. *PLoS Comput. Biol.* 15, e1007422 (2019).

51. Varlet, M. et al. Impairments of social motor coordination in schizophrenia. *PLoS One* 7, e29772 (2012).

52. Marsh, K. L. et al. Autism and social disconnection in interpersonal rocking. *Front. Integr. Neurosci.* 7, 4 (2013).

53. Dermody, N. et al. Uncovering the Neural Bases of Cognitive and Affective Empathy Deficits in Alzheimer’s Disease and the Behavioral-Variant of Frontotemporal Dementia. *J. Alzheimers Dis.* 53, 801–816 (2016).

54. Oldfield, R. C. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113 (1971).

55. Davis, M. H. & Others. A multidimensional approach to individual differences in empathy. (1980).

56. Afsar, O., Tirnakli, U. & Marwan, N. Recurrence Quantification Analysis at work: Quasi-periodicity based interpretation of gait force profiles for patients with Parkinson disease. *Sci. Rep.* 8, 9102 (2018).

57. Fraser, A. M. & Swinney, H. L. Independent coordinates for strange attractors from mutual information. *Phys. Rev. A Gen. Phys.* 33, 1134–1140 (1986).

58. Bradley, E. & Kantz, H. Nonlinear time-series analysis revisited. *Chaos* 25, 097610 (2015).

59. Rulkov, N. F., Sushchik, M. M., Tsimring, L. S. & Abarbanel, H. D. Generalized synchronization of chaos in directionally coupled chaotic systems. *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics* 51, 980–994 (1995).

60. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models using Lme4. *arXiv [stat.CO]* (2014).
Legends

**Figure 1. Experimental design.** The study was designed in a Modality (Visual, Auditory) x Coupling (Coupled, Uncoupled) factorial structure, resulting in the following conditions.
1. **Visually Coupled.** Participants tapped along with an auditory metronome, while looking at the partner's hand tapping. The view of their own hand was hidden by a screen placed on the table. They were explicitly asked to neglect the partner's movements and focus on following their own metronome.
2. **Visually Uncoupled (control).** Participants tapped along with an auditory metronome, while looking at their own hand. The view of the partner's hand is hidden by a screen placed on the table.
3. **Audiotorily Coupled.** Participants tapped along with a flickering LED, while hearing the sonification of the partner's tapping. The view of all the hands was hidden with screens placed on the table. They were informed that the sounds they will ear are produced by the partner, and were explicitly asked to neglect them and focus on following their own metronome.
4. **Audiotorily Uncoupled (control).** Participants tapped along with a flickering LED, while hearing the sonification of their own tapping. The view of all the hands was hidden with screens placed on the table.

**Figure 2. Analysis pipeline.** The procedure is here presented as applied to the timeseries generated by the metronomes' onsets, showing dyadic behavior under the ideal condition of perfect compliance with the task. 
A) Starting from discrete onsets, B) we interpolated each timeseries with a sine function to model the behavior as a system of coupled oscillators \(^50\) and C) separately embed them in their respective phase-spaces for JRP\(s\) computation \(^20\). From the representation of one full cycle of the metronomes' behavior, it is evident how the two oscillators smoothly de-phase and get back in-phase at the end of the cycle. 
D) JRP\(s\) were computed of each individual cycle and E) summed up, returning a 2-D matrix of the density of recurrence points over the trial structure. Finally, by looping over the column and summing all the rows of the matrix, F) we compressed the representation into a 1-D timeseries of the total count of recurrences over the cycle structure. 

**Figure 3. Recurrence score over average cycle.** The sum of 10 JRP\(s\) from Visual Coupling and Visual Control conditions are compressed to a 1-D recurrence score timeseries. It is evident how visual coupling modulates the density of recurrence point and its distribution over the average cycle in this representative dyad (#5). Compared to the control condition, the recurrence score was higher on average and exhibited local maxima at the extremities of the cycle (i.e., in proximity of the in-phase attractor point).

**Figure 4. Orthogonal polynomials.** The time-varying dimension of a response variable over the metronomes cycle can be modelled with higher-order polynomials. In our model, intercepts correspond to the overall average of the recurrence score: the main effects of categorical predictors.
indicate global differences across experimental conditions independently from the temporal profile of the response variable. On the other hand, effects of categorical predictors on the polynomial terms indicate the local effect of attractor points on the $x$-axis. For the linear component (1st order), the parameter estimate corresponds to the slope of the line; for the quadratic component (2nd order), it corresponds to the parabolic curvature; for the cubic and quartic components (3rd and 4th order), it corresponds to the sharpness of the peaks on the inflexion points. The order of the polynomial is ideally chosen based on hypothesis and on the nature of the data, is should be confirmed by the empirical data, and should be allow a straightforward interpretation of the effects.

In our case, orders from 2 to 4 seem suited candidates for modelling different possible attractor landscapes, given the symmetric measure of the cycle and inflection points around in-phase and anti-phase regions. A full model based on orthogonal polynomials includes all terms up to the chosen order.

Figure 5. Coupling strength as a function of the attractor landscape. A) Recurrence score (on the log-scale) over the average of 10 cycles, for 14 dyads (colored dots). Grand-average and error bars are represented in black. The reader can visually appreciate the consistency of the "seagull-shaped" pattern in condition of visual coupling and the flat line in the respective control condition. The condition of auditory coupling shows a recurrence score higher on average but no clear pattern over the average cycle, and overall higher variability. B) Recurrence score predicted by our 2nd order polynomial model. The inflection of the parabolic term and the negative slope of the linear term successfully capture the influence of the in-phase attractor points in the visually coupled condition. Arguably, this is the most parsimonious model for our data in the framework of growth curve analysis. C) Recurrence score predicted by our 4th order polynomial model. The quartic fit manages to capture more fine-grained modulations, particularly the local maxima on the inflexion points at the extremes of the cycle. The profile recalls the typical 'seagull-effect' originally reported in and discussed in in the context of bimanual coordination dynamics. However, for the sake of interpretation of parameter estimates, we opted for the simplest possible model: given the absence of a local maximum on the anti-phase point, a single parabolic inflexion point was sufficient to explain the effect of visual coupling. D) Within-dyads comparison of the recurrence score over the average trial, across experimental conditions. Dots represent the empirical data, whereas continuous lines represent the full model (up to the 2nd order polynomial) fitted to the data.

Table 1. Model summary. All significant effects are marked with an asterisk, and the associated $p$-values are highlighted in bold. We reported a significant main effect of Modality and significant 3-way interactions with linear and quadratic functions of Time. In the framework of orthogonal polynomials, coefficient estimates represent a measure of effect size: the intercept represents the average increase in the response variable contrasted to the baseline (i.e., the Uncoupled level), the slope of the linear term represents the change in the response variable per unit of time, and the quadratic term represents the steepness of the parabolic inflexion. Given the response variable was transformed to the log-scale, the natural exponential of the coefficient estimates should be used for interpretation purposes. The significant correlation with the quadratic function of time was not discussed in the text, for it does not have relevance in the context of the experimental design.