

D2.6

Results on prediction in complex action execution and observation - Phase II

Project No	GA824160
Project Acronym	EnTimeMent
Project full title	ENtrainment & synchronization at multiple TIME scales in the MENTal foundations of expressive gesture
Instrument	FET Proactive
Type of action	RIA
Start Date of project	1 January 2019
Duration	48 months

Table of Contents

Introduction	4
Executive summary.....	4
2.3.1 Orchestra violin sections and conductor	7
Figure 1: Schematic representation of the main results for intra-group analysis.	10
Figure 2: Schematic representation of the main results for inter-group Granger-Causality analysis across the conductor and the two sections of violinists	11
References.....	11
2.3.2 Dance to Sync.....	13
2.3.2.1 Phase I – Synchronization and perceptual interruption.....	13
Figure 1: Four topologies during familiar human group cooperation situations.....	15
Figure 2: Main results of Experiment 1.	16
Figure 3: Main results of Experiment 2.	17
References.....	18
2.3.2.2. Phase II – Leadership patterns during synchronization in complex human networks....	18
Figure 1: Description of the leadership scenarios.	20
Figure 2: Distribution and characterization of the leadership scenarios.....	22
Figure 3: Distribution and characterization of the leadership scenarios in absence of social interaction.....	23
2.3.4 Time to Sync	27
2.3.4.1 Phase I - EmoSync: human multi-scale group synchronisation	27
Figure 1: The experimental setup.	28
Figure 2: The Order Parameter extraction visualization.	29
Figure 3: Physiological and behavioural synchronisation metrics summary.	30
2.3.3.2 'Spook and play': What are the effects of emotional induction/perturbation using affective sounds from IADS-2 battery during playing a disembodied, multiplayer mirror game?	32
Figure 1: Experimental setup.	34
Figure 2: Trial construction	35
Figure 3: Multi-level synchronisation analysis example.....	36
Figure 4: Mean order parameter per number of negative emotional inductions in the group.	38
Figure 5: Number of cycles per positive inductions in the top panel per group, and negative bottom panel per group.	39
Figure 6: Average number of cycles per emotional induction condition.....	40

Figure 7: Individual synchronization index values for each emotional condition	41
Figure 8: Distribution of the Individual Synchronization Indexes per emotional condition.	42
2.3.4 The sensitivity of intrapersonal coordination to interpersonal coupling	45
Figure 1: Coupling between violin sections and composer.....	46
Figure 2: Average cross-correlation functions between head and bow motion	50
Figure 3: Violin plots of the amount of covered space and spatial dispersion of bow and head positions in each experimental condition.	51
Table 1: Mean inter-distance between positional datapoint and convex hull volume indicating head and bow trajectories.....	51
Figure 4: Power spectral density of the motion of the bow and the head of V1 for each experimental condition.....	52
Table 2: Power values extracted from the power spectral density spectra at the 4 frequency bins that corresponded to musical metrical levels compared across experimental conditions for the motion of the bow.	53
Table 3: Power values extracted from the power spectral density spectra at the 4 frequency bins that corresponded to musical metrical levels, compared across experimental conditions for the motion of the head.....	53
Table 4: Correlation coefficients representing the co-evolution of V1's head and bow power spectral density at the 4 frequency bins compared across experimental conditions.....	54
Figure 5: Power spectral density of the motion of the bow and the head of V2 for each experimental condition.....	55
Table 5: Correlation coefficients representing the co-evolution of V1's and V2's head and V1's head and V2's bow power spectral density at the 4 frequency bins compared across experimental conditions.....	56
Table 6: Vector length mean relative angle and mean absolute angle of the relative phase between head and bow motion, computed at each metrical timescale compared across experimental conditions.....	57
2.3.5 Tracking the leader: gaze behaviour in group interactions.....	58
Figure 1: Depiction of gaze estimation methods applied in the study and result matrix.	59
Table 1	60
References.....	61
2.3.6* Capturing human movement and shape information from small groups to extract expressive and social features - using marker-less techniques	62
Figure 1: Image from a musical piece composed by Johannes Brahms.....	64
Figure 2: Measure output for Position x Texture	65
Figure 3: Measure output for Texture x Pair.....	66
Figure 4: Measure output for Texture x Position x Pair	66

Table 1: Mean and standard deviation of estimates of pulse clarity and event density as a function of texture (homophonic and polyphonic) and phrase position.....	68
References	68

Introduction

This deliverable reports on the progress on the research conducted between M7-M36 of the EnTimeMent project with regards to the complex action execution and observation axis, with particular interest in group motor behaviour (involving three people or more, $n > 2$). This part of the project is fed by the research and theoretical work developed in the Phase I and Phase II of the EnTimeMent project, focused on the individual ($n=1$) and dyadic studies ($n=2$), which are reported in D2.1, D2.2; and D2.3, D2.4 respectively, but focuses on complex action of multiple agents interacting with each other.

The numbering of the studies reported herein, refers to the last version of deliverable *D1.2 Research Requirements*, which provided an overview on the methodological background and know-how of the EnTimeMent studies. In comparison to the previous version of this deliverable (D2.5), all experiments reported herein have their data collection finalised, processed and analysed. Importantly, data originally collected in 2.3.1 (*Orchestra violin sections and conductor*) has been completely re-processed and analysed to start answering research questions specified in D1.2 as in 2.3.4. The Experiment 2.3.6 as added as supplementary to the original list in D1.2.

Executive summary

Our brains have been carved by evolution to act together with other people, towards long-term mutual goals. Studies reported below investigate how people combine sensory information to achieve coordination with others, drawing in parallel from their internal mental representations (i.e., memory), adaptation and prediction processes. First project, **2.3.1** at IIT-FE, harnessed a naturalistic interaction in the violin ensemble to understand those sensorimotor dependencies. Two channels of communication were explored (head and bow movements of the violin players - section one and two; in relation to the conductor) under full and perturbed access to the visual information of other players. Study has found that bow kinematics in players in two violinists' sections exhibit a robust leader-follower relationship with the conductor. Study showed that bow kinematics resist visual perturbation (part of the section being deprived of the vision of the other part) due to reliance on memory of the score and aforementioned primary role of flow of information

from conductor in leading the group performance. For the head movement of players, different pattern emerged revealing emergence of strategy to maximize performance accuracy in the section deprived of visual contact with conductor by tighter coordination of within the section. In parallel, a drop in intra-group coordination was noted for the section with preserved visual contact with the conductor, but becoming a coordination hub for the section deprived of visual information and conductor. This emergence of strategy demonstrated the ability of experts, such as violin players, to adapt intra- and inter-group dynamics to maintain robust coordination in a context of a group perceptuomotor performance. In the last year of the project, a new approach of analyses was carried out to investigate intra-body coupling between head and bow movements. These preliminary analyses (reported as **2.3.4** here) show that overall intrapersonal coordination of head and bow motion was increased due to the visual perturbation. Interestingly, head (but not bow) movements show increased spectral complexity, to reflect more closely the metrical structure of the score. These analyses will continue in the final year of the project to describe the specific strategies implemented by expert musicians to cope with a visual perturbation via the adaptation of intra-body movement coordination. The role of the perceptual contact intra-group and with the leader, was dissected in the **2.3.5** project looking into the eye contact (gaze) relationship during group discussion with manipulation of the factors of the leadership style (authoritarian versus democratic, and high and low time pressure condition). Leaders looked less at others and, conversely, were looked at more as compared with followers. Also, leaders were involved in and caused more episodes of mutual engagement, relative to followers. This study has demonstrated the key contribution of the gaze exchange in social interactions, building up on the previous results from **2.3.1** and **2.3.2**. Another future direction of this research is the study of social information transmission in groups and its relationship to leadership. IIT Genoa group has extensive experience with methods for quantifying the encoding and readout of information in movement kinematics (see *2.1.17 Intersecting action and perception in autism spectrum disorders at the single-trial level in D3.6 and Patri et al., 2020*). We currently plan to apply these methods to track the transmission of social information in dyads and groups.

Next project, based at EuroMov, **2.3.2 (.1 Phase I)** continued to further explore how humans (expert- dancers; and non-expert - sport students) maintain regimes of coordination despite a transient loss of visual coupling during pendulum-based synchronization task. This study has strengthened the idea that the ‘topology’ - spatial alignment between people mediates access to the real-time, perceptual information about movement of others plays a key role in the succeeding in behavioural cohesion. Participants were able to maintain synchronization of swinging the pendulum with matching frequencies up to 7 seconds after the loss of perceptual contact (manipulation of field of view with visual occlusion glasses), with more accurate performance observed

for experts, revealing both short timescale memory constraints and role of perceptuo-motor expertise (long timescale) in achieving coordination success. In addition, a follow up study (**2.3.2.2. Phase II**) conducted with external collaborators on leadership during group synchronization, compared two types of leadership: Phase leadership and influence leadership and is summarised herein.

Two **2.3.3** *Time to sync* studies (EMOSYNC and Play'n'Spook) reported hereunder, shed light into the impact of emotion on the ability to synchronize movements with others. We reported in **2.3.3.1** EMOSYNC favourable effect of positive emotional induction on the synchronisation metrics during improvisation mirror game task for triads in comparison to emotionally neutral condition. In the final year of the EnTimeMent project, as a continuation of EMOSYNC, we will restrict the improvisational movement aspect, which was characterized by a large inter-individual variance, to a more controlled experimental paradigm of an air finger tapping. In addition, the impact of movement on the heart rate in the first experiment was overwhelming. We could not systematically discriminate, without greatly increasing the experiment duration, what changes were produced in different experimental emotional conditions. Rather than affective reactions to stimuli, the movement requires considerably larger adjustments of a heart. That is why in the follow-up experiment, planned for the remaining part of the EnTimeMent project, the participants will be seated with comfortably placed elbow on the table and the only movement executed will be finger oscillations in the air (embodied synchronisation). In the **2.3.3.2** Play'n'Spook study we demonstrated that group emotion has impact on the ability to synchronise in a digital, disembodied synchronisation scenario with four participants. In follow-up study that is ready to launch at EuroMov we will run experiment on passing an emotion-laden object in a cooperation type of task for a small group. The goal is further investigation of emotion embodiment and propagation via movement in multi-agent scenario and whether emotion can propagate via movement of participants, without visual access to facial expression of participants.

A final study included in this report, **2.3.6** (delivered by UNIGE, but not previously specified in D1.2) investigated the relationship between the interpersonal coordination of body motion and musical structure (in a musical ensemble using marker-less motion capture techniques), with a practical methodological advance of proposing an alternative technique for the automated analysis of human body movements. This method has been applied specifically to computing dyadic synchronization between co-performers in musical ensembles, and provides a broad perspective with a kinematic and audio features analysis. We report that interpersonal coupling tends to be stronger between performers for polyphonic than homophonic textures. In the final year of this project, further work will include analysis of those features in the context of multiple timescales.

Taken together, Phase I and II results reported in this deliverable, pushed forward the state-of-the-art models of human coordination in complex group situations involving both action execution and observation. We addressed multiple gaps in the body of research and emphasised importance of multiple timescales and modalities approach (movement, gaze and cardiac activity data signals) and the need for inclusion of emotion in analysis of human sensorimotor group behaviour.

2.3.1 Orchestra violin sections and conductor

For a full description please see: Hilt P.M., Badino L., D'Ausilio A., Volpe G., Fadiga L., Camurri A. (2019) Multi-layer adaptation of group coordination in musical ensembles. *Sci Rep*, 9: 5854.

Successful human-to-human interaction requires important behavioral adaptation, as well as prediction. A large body of literature has focused on cooperation towards shared goals, where humans must combine available sensory information with internal movement production models (Wolpert et al., 2013; Sebanz et al., 2009; Jeannerod et al. 2001; Friston et al. 2011). In this regard, researchers investigated how dyads achieve interpersonal simple sensorimotor coordination, such as walking side-by-side (van Ulzen et al, 2015) or rocking in rocking-chairs (Richardson et al. 2007). In such contexts, co-actors continuously influence each other and tend to spatially and temporally synchronize their movements. Beside imitation, action complementarity plays a key role in inter-individual coordination with the goal of achieving efficient collaboration (Newman-Norlund et al., 2007). Social interaction indeed goes beyond synchronization with other's actions and relies also on inferring others' motor goals and intentions to generate a context-appropriate action. To achieve fast inter-individual coordination, individuals may build internal predictive models of other's behavior. In function of the context, the most appropriate motor model is compared with the current observed movement, to generate a prediction error (Friston et al., 2011) and update own motor planning (Sebanz et al., 2006). Due to the technical and analytical complexity in exploring the details of human sensorimotor interaction, only few experiments went further than a dyadic set-up (Fessler et al., 2016; Dikker et al., 2017, Alderiso et al., 2016). However, in daily life, things are usually much more complex. For instance, during a conversation, information is sampled through multiple channels (e.g., vision, audition), sometimes in parallel (e.g., information in the foreground and information from the background) and at different temporo-spatial scales (e.g. slow whole-body movements versus fast lip motions). At the same time, different kinds of information may be conveyed in parallel through different channels. For

example, in speech, bodily gestures and spoken words are generally co-expressive (McNeill, 2000). On this basis, communication requires flexible means to integrate multimodal data, across multiple timescales and act accordingly. Therefore, proper quantification of (realistic) group coordination is today one of the key missing elements to understand how humans manage to interact with others by efficiently selecting, processing and sending information.

In this context, ensemble musicians have been proposed as an ideal model, by keeping the key multidimensional properties of natural sensorimotor interaction, but allowing relatively good experimental control (Volpe et al., 2016; D'Ausilio et al. 2015). Beyond global descriptions of musician's pattern of relationships, the complexity of these kinds of scenarios could also be exploited to distinguish and evaluate the existence of multiple channels of communication as well as their respective role in efficient coordination. In previous studies, one representative kinematic parameter was used to extract global coordination (D'Ausilio et al., 2012; Badino et al., 2014, Chang et al., 2017., Chang et al., 2019). However, we know that movements of different body parts may convey substantially different types of information. For instance, bow movements in violinists directly control the sound output (i.e., instrumental gestures), whereas complementary torso oscillations may serve a secondary communicative purpose (ancillary gestures (D'Ausilio et al., 2015)). More importantly, movements of different body parts may act as different channels of communication, possibly with different roles depending on the specific communication mode. For example, within a quartet (Badino et al., 2014; Chang et al., 2017; 2019), musicians have specific roles while in orchestras, musicians generally play in distinct sections (e.g., sections of violinists). This means that in the orchestra scenario, different modes of communication coexist: a complementary coordination with the conductor and other musicians, in parallel with an imitative coordination with musicians of the same group (playing the same score).

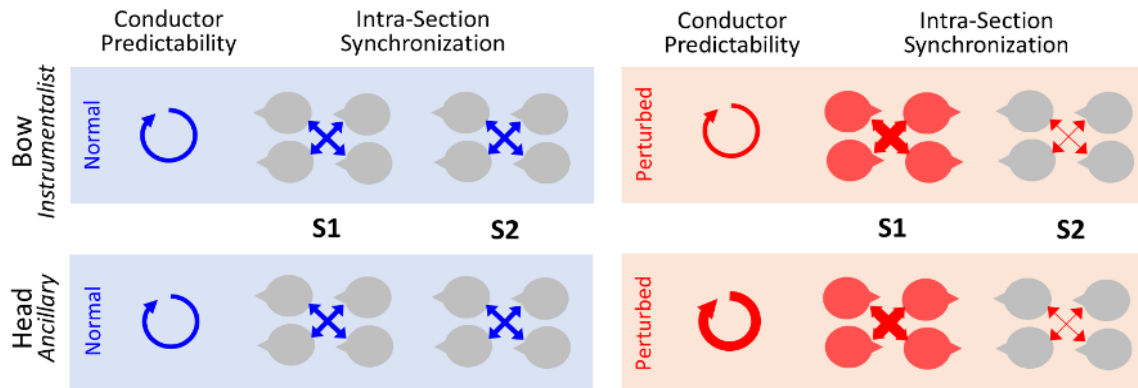
In the present study, we aim at answering two scientific questions: whether different channels of communication exist and whether they carry different information across modes of communication. We had a chamber orchestra playing music while we recorded bow and head kinematics (instrumental and ancillary movements) of a first and second section of violinists (four violinists in each section) as well as the arm and head kinematics of two different conductors. In one experimental condition we applied a perturbation to the orchestra sensorimotor information flow. The perturbation consisted in half-turn rotation of the first section of violinists so that they faced the second section and couldn't see the conductor anymore. This perturbation modifies the perceptuo-motor context of the first section of violinists, placing also the second section and the conductor into a novel playing situation. By doing so, we analyzed inter-group complementary coordination as well as intra-group imitative coordination (modes of communication), through different

channels of communication (instrumental and ancillary movements) during different playing situations (normal and perturbed).

Our results demonstrate that the pattern of sensorimotor information carried by two selected movements (head and bow) are distinct. Bow kinematics exhibit a robust leader-follower relationship between the conductor and the two violinists' sections. This pattern is substantially not affected by the experimental manipulation of the sensorimotor information flow (perturbed condition) except for a decrease in communication between the first section and the conductor. The fact that the perturbation did not dramatically alter the information exchanged via instrumental movements suggests an important role of memory, score reading and residual sensory cues. Indeed, musicians train for several hours and may rely on rehearsal memory to cope with the perturbation, at least for what concerns pure instrumental execution. At the same time, there is also a clear directionality of the information flow from conductor to musicians, which confirms the idea of a predominant role of the conductor in group management. Ancillary movements, instead, are supposed to convey slower frequency signals possibly related to the expressive component of musical execution, which is more likely to be affected by perturbation of the interaction dynamics. In fact, in head data, the perturbation produced clear alteration of the communication pattern. Communication between the first section and the conductor or the second section was reduced. At the same time, communication between the second section and the conductor increased in both directions. This global increase suggests a greater need for information exchange during the perturbation and considering that conductors and S2 did not change their positions, we observe a quantitative but not a qualitative alteration of their communication. Instead, moving to the relationship between S1 and S2 we observe a complete reversal of their mutual communication. Before the perturbation, the first section provided larger causal drive towards the second section, while after, the second section led the first. During the perturbation, the first section no longer had visual contact with the conductor, significantly reducing his role in leading orchestra dynamics. Even if we cannot exclude the contribution of haptic, acoustic or residual visual information, the remaining influence that the conductor exert on S1, seems to be mediated by the new role played by S2. Although violinists of the second section did not actually change their position, they are the only ones establishing direct face-to-face communication with both the first section and conductor. Interestingly, they seem to increase their normal communication with conductors, while at the same time they dramatically change the way they communicate with S1. Correspondingly, our results suggest that S2 musicians were implicitly invested with far more centrality in orchestra coordination dynamics.

The distinct modulation of head versus bow kinematic parameters provides a demonstration of the multi-level complexity of musicians' coordination. At the same time,

another important aspect is related to the co-regulation of different modes of interaction. In our experimental context, each violinist must exchange information with other musicians of the same section (playing the same musical score – intra-group imitative coordination – Figure 1) and with other participants (playing different parts – inter-group complementary coordination – Figure 2). We used PCA to complement inter-group Gca



analysis with an estimation of intra-section imitative coordination. Both kinematic parameters highlighted similar patterns of results. Due to the lack of communication with the conductor, the first section became more coordinated, in the probable attempt to maximize performance accuracy. On the contrary, the second section that was endowed with the central role of being the communication hub, reduced intra-group coordination. This may be driven by a need to gain the necessary degrees of freedom to lead communication with S1 and be the sole interlocutor of the conductor. Therefore, here we show that to modulate inter-group dynamics, S2 violinists had to penalize imitative coordination at the intra-group level.

Figure 1: Schematic representation of the main results for intra-group analysis. Results associated with the two channels, bow and head, are displayed respectively in the upper and lower panel. Circular arrows displayed in the left panel represent the strength of conductor predictability (ARfit). Middle and right panels represent the intra-group synchronization (%PC1) for both sections of violinists. Thickness of the arrow represents the strength of the effect in each experimental condition: normal (blue) and perturbed (red).

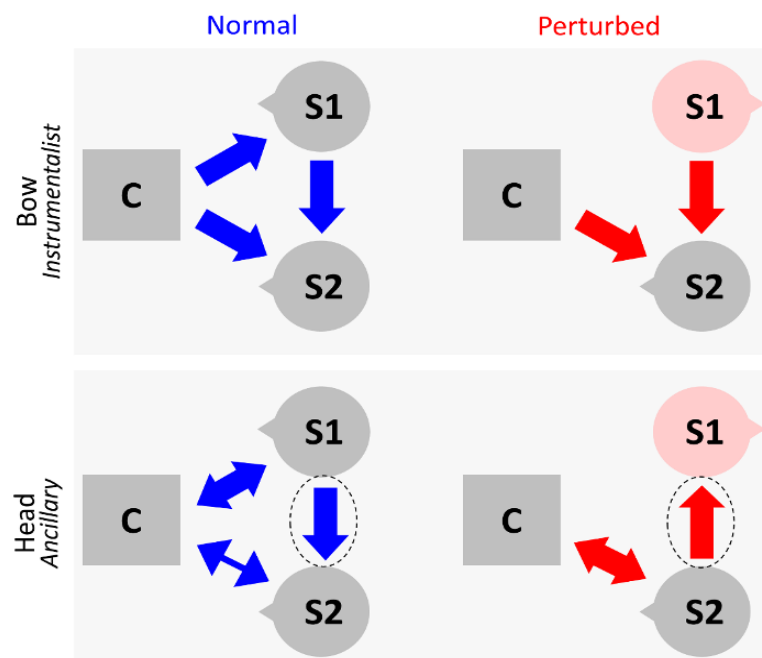


Figure 2: Schematic representation of the main results for inter-group Granger-Causality analysis (i.e. inter-group coordination) across the conductor (C) and the two sections of violinists (S1 and S2). Results associated with the two channels, bow and head, are displayed respectively in the upper and lower panel. Directional arrows illustrate inter-group coordination (C, S1 and S2), in the normal (blue) and perturbed (red) condition. Arrows thickness represents the interaction's strength. A bidirectional arrow indicates similar gca values for the two directions (i.e. group 1 G-causes group 2, as much as group 2 G-causes group 1). On the opposite, a unidirectional arrow indicates the direction of the larger gca value (e.g. group 1 G-causes more group 2, than the inverse). To highlight the difference between the two conditions, we did not represent the arrow between C and S1.

References

- Alderisio, F., Fiore, G., Salesse, R. N., Bardy, B. G. & di Bernardo, M. (2016) Interaction patterns and individual dynamics shape the way we move in synchrony. arXiv. at <<http://arxiv.org/abs/1607.02175>>
- Badino, L., D'Ausilio, A., Glowinski, D., Camurri, A. & Fadiga, L. (2014). Sensorimotor communication in professional quartets. *Neuropsychologia* 55, 98–104.
- Chang, A., Livingstone, S. R., Bosnyak, D. J. & Trainor, L. J. (2017) Body sway reflects leadership in joint music performance. *Proc. Natl. Acad. Sci.* 114, E4134–E4141.
- Chang, A., Kragness, H. E., Livingstone, S. R., Bosnyak, D. J. & Trainor, L. J. (2019) Body sway reflects joint emotional expression in music ensemble performance. *Sci. Rep.* 9.
- Colley, I. D., Varlet, M., MacRitchie, J. & Keller, P. E. (2018) The influence of visual cues on temporal anticipation and movement synchronization with musical sequences. *Acta Psychol. (Amst)*. 191, 190–200.
- Codrons, E., Bernardi, N. F., Vandoni, M. & Bernardi, L. (2014) Spontaneous group synchronization of movements and respiratory rhythms. *PLoS One* 9.
- D'Ausilio, A., Novembre, G., Fadiga, L. & Keller, P. E. (2015). What can music tell us about social interaction? *Trends Cogn. Sci.* 19, 1–4.
- D'Ausilio, A. et al. (2012). Leadership in orchestra emerges from the causal relationships of movement kinematics. *PLoS One* 7.
- Dikker, S. et al. (2017). Brain-to-Brain Synchrony Tracks Real-World Dynamic Group Interactions in the Classroom. *Curr. Biol.* 27, 1375–1380.
- Fessler, D. M. T. & Holbrook, C. (2016). Synchronized behavior increases assessments of the formidability and cohesion of coalitions. *Evol. Hum. Behav.* 37, 502–509.
- Friston, K. J., Mattout, J. & Kilner, J. M. (2011). Action understanding and active inference. *Biol. Cybern.* 104, 137–160.
- Jeannerod, M. (2001). Neural simulation of action: a unifying mechanism for motor cognition. *Neuroimage* 14, S103–S109.

McNeill, D. *Language and Gesture*. (2000).

Newman-Norlund, R. D., Noordzij, M. L., Meulenbroek, R. & Bekkering, H. (2007). Exploring the brain basis of joint action: Co-ordination of actions, goals and intentions. *Soc. Neurosci.* 2.

Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R. L. & Schmidt, R. C. (2007). Rocking together: Dynamics of intentional and unintentional interpersonal coordination. *Hum. Mov. Sci.* 26, 867–891.

Sebanz, N., Bekkering, H. & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends Cogn. Sci.* 10, 70–76.

Sebanz, N. & Knoblich, G. (2009). Prediction in Joint Action: What, When, and Where. *Top. Cogn. Sci.* 1, 353–367.

van Ulzen, N. R., Lamothe, C. J. C., Daffertshofer, A., Semin, G. R. & Beek, P. J. (2008). Characteristics of instructed and uninstructed interpersonal coordination while walking side-by-side. *Neurosci. Lett.* 432, 88–93.

Volpe, G., D'Ausilio, A., Badino, L., Camurri, A. & Fadiga, L. (2016). Measuring social interaction in music ensembles. *Philos. Trans. B* 371, 20150377.

Wolpert, D. M., Doya, K. & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philos. Trans. R. Soc. London* 358, 593–602.

2.3.2 Dance to Sync

2.3.2.1 Phase I – Synchronization and perceptual interruption

For a full description please see: Bardy, B. G., Calabrese, C., de Lellis, P., Bourgeaud, S., Colomer, C., Pla, S., & di Bernardo, M. (2020). Moving in unison after perceptual interruption. *Scientific Reports*, 10:18032 doi.org/10.1038/s41598-020-74914-z.

Humans and other animals often cooperate in small or large ensembles, for anti-predation, for producing a collective performance, or sometimes just for entertainment. Among all sorts of cooperative behaviours, synchronization in space and/or in time of the members of the group is particularly present in the human repertoire. It is often rooted in perceptuo-motor synergies in which proximal (e.g., postures, breaths) or distal (e.g., gazes, voices, hands, legs) parts of the body are delicately locked, for brief or long periods of time, in frequency and in phase (Kelso et al., 1986; Strogatz et al., 2004). In sport, music and dance moving in unison is either the goal or clearly contributes to it, and results from both (i) personalized characteristics and (ii) the way individuals are coupled together. Synchronization also requires a coupling function between the various systems involved, whether these systems are of physical, biological, or social origin (Damm et al., 2019). Perceptual contact is the most natural form of coupling between agents in a group.

Of interest for the present research is the recent discovery that certain topologies of the spatial organization of members in the group affect the strength and symmetry of perceptual coupling. The current study targeted the powerful capacity of humans to maintain regimes of synchronization despite a transient loss of perceptual (i.e., visual) coupling as it often occurs in daily scenarios. This phenomenon occurs for instance when a group of people continue to walk at the same pace even after they separate, or when dancers in a choreographic performance maintain body synchronization during a transient lack of visual connection. This capacity is a solid contributor to a wide range of social performances, in sport or at work. It relies on our practical ability to internalise previously-produced movement patterns in a social context, and to maintain them when alone for a certain amount of time. The aetiology of this ability is somewhat dual (physical and neural). One approach, the individual memory approach, considers this persistence effect as a witness of our capacity to prolong a movement pattern previously produced under a certain goal (intentional group synchronization) into a new context (solo action). In contrast, the social memory approach (Oulier et al., 2018; Nordham et al., 2018) suggests that persistence after visual interruption is the consequence of the mental simulation of the social interaction previously created. It also predicts that a certain proximity in

individual movement frequencies fastens the route toward synchronization and helps to maintain a stronger coordinate regime during perceptual coupling and temporarily after its interruption but anchors these evidences into the social benefits of synchronization. Here we investigated the dynamics of voluntary synchronization, in groups composed of seven participants, manipulating their similarity, spatial organization and the presence or duration of visual coupling. Participants were engaged in an intentional group synchronization task and had to swing a pendulum in order to achieve unison in space and in time (phase synchronization).

This task was selected as (i) it is extremely easy to learn and perform, (ii) it has been documented before in a dyadic context (Schmidt et al., 1990), and (iii) it allows a simple yet precise control over each participant's natural frequency. Each trial started with an eyes-closed period, in which each player oscillated their own pendulum at their preferred pace. This was followed by an eyes-open period, where they had to reach synchronization as fast as possible. The last period was again an eyes-closed sequence, split into two-time intervals of equal length, in order to better identify the possible presence and duration of a memory effect.

We have conducted two experiments in this study. In Experiment 1, participants' similarity was controlled by manipulating the pendulum's inertia and hence the natural frequency of the players' oscillatory motion. This enabled us to evaluate the influence of the players' similarity and graph structure on the emergence and quality of group synchronization. Specifically, four conditions were considered, involving (i) individual oscillations (solo), and three collective oscillations (ii) at the same shared frequency (all matched), (iii) at the same frequency for six out of the seven players (all matched but one), and (iv) at seven different frequencies corresponding to each player's preferred pace (natural). Each group performed the experiments in four different interaction patterns among players (i.e., topologies), implemented through the combination of the spatial location of each participant and the use of home-made goggles limiting the field of vision to the desired location.

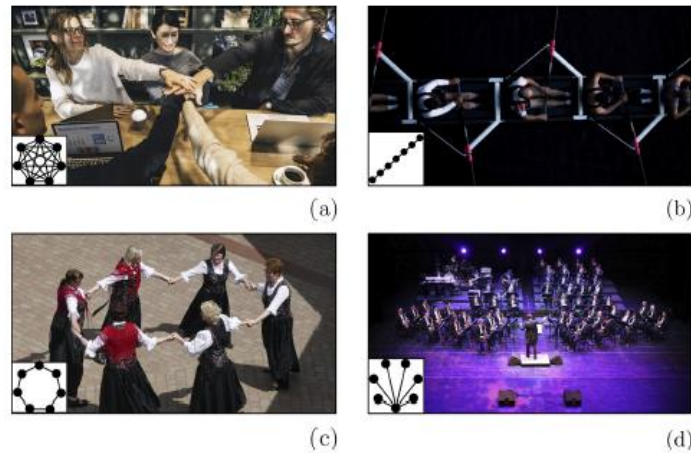


Figure 1: Four topologies during familiar human group cooperation situations, with various coupling modalities. (a) Complete graph: an ordinary organization during everyday working meetings; (b) Path graph: often present in sports, for instance in team rowing where partners are mechanically and visually coupled to two neighbours, except for the first and last rowers; (c) Ring graph: a common structure in many popular dances or among children at play (round dance); (d) Star graph: typical of musical ensembles, for instance when orchestra members are visually coupled only to the director.

Namely, the four topologies were Complete graph, Path graph, Ring graph and Star graph. In Experiment 2, homogeneity among the players was manipulated at a different scale, by comparing groups of novices with groups of certified dancers. For homogeneity, we predicted that similarity would strengthen synchronization, irrespective of graph topology (Experiment 1, see Figure 1), and that dancers would maintain a more solid synchronization regime compared to non-dancers (Experiment 2). For topology, we expected that Complete and Star graphs, that were observed to maximize synchronization metrics during visual contact, would still be associated with higher levels of coordination after visual interruption. Furthermore, we predicted that a stronger memory effect would be present in the case of higher homogeneity between participants (similar pendulum frequencies in Experiment 1 and dancers in Experiment 2) and in graphs producing higher perceptual exchanges (Complete and Star graphs). Experimental results are summarised on Figure 2 and 3.

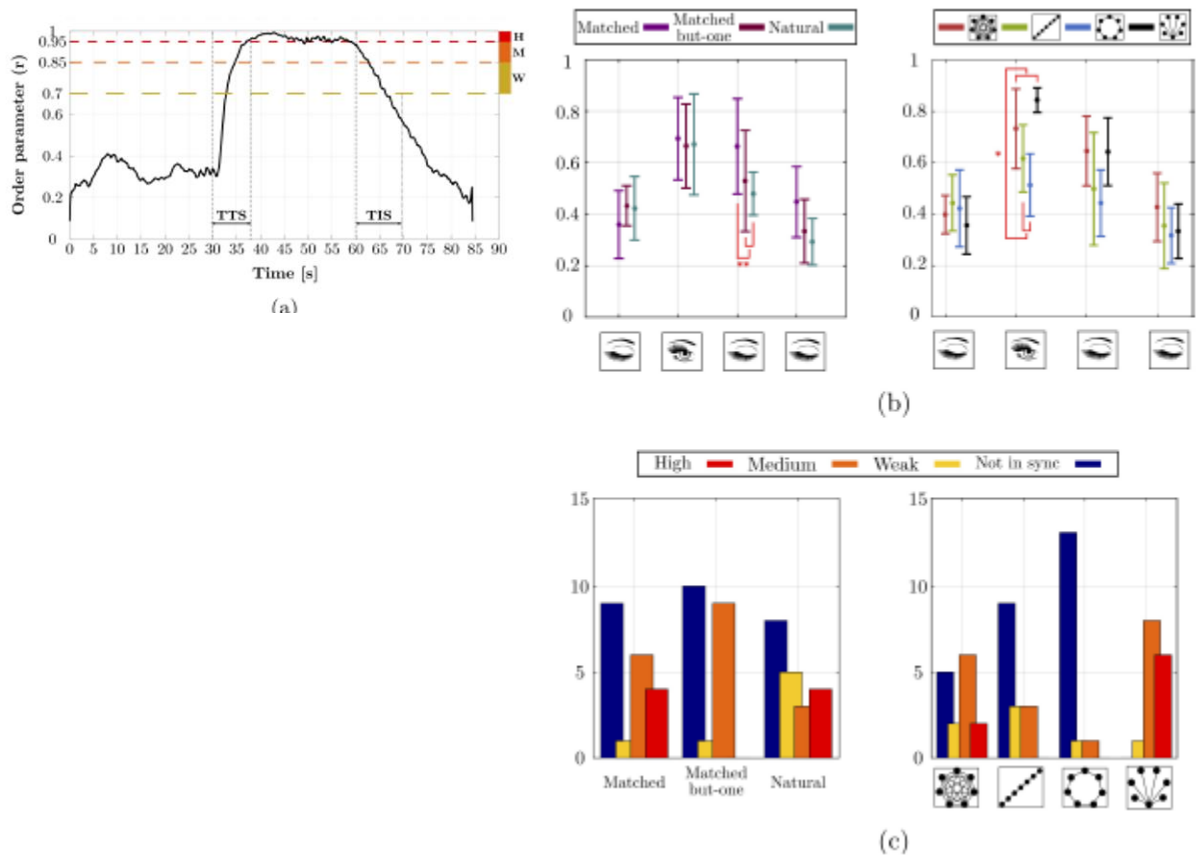


Figure 2: Main results of Experiment (a) a representative example of phase synchronization r across periods of absence and presence of visual coupling (Time To Sync TTS = 7:49 s and Time In Sync TIS = 9:78 s); (b) mean and standard deviation of phase synchronization across homogeneity (left panel) and topology conditions (right panel), $n = 240$; (c) distribution of phase synchronization levels for Similarity (left panel) and Topology (right panel), $n = 60$.

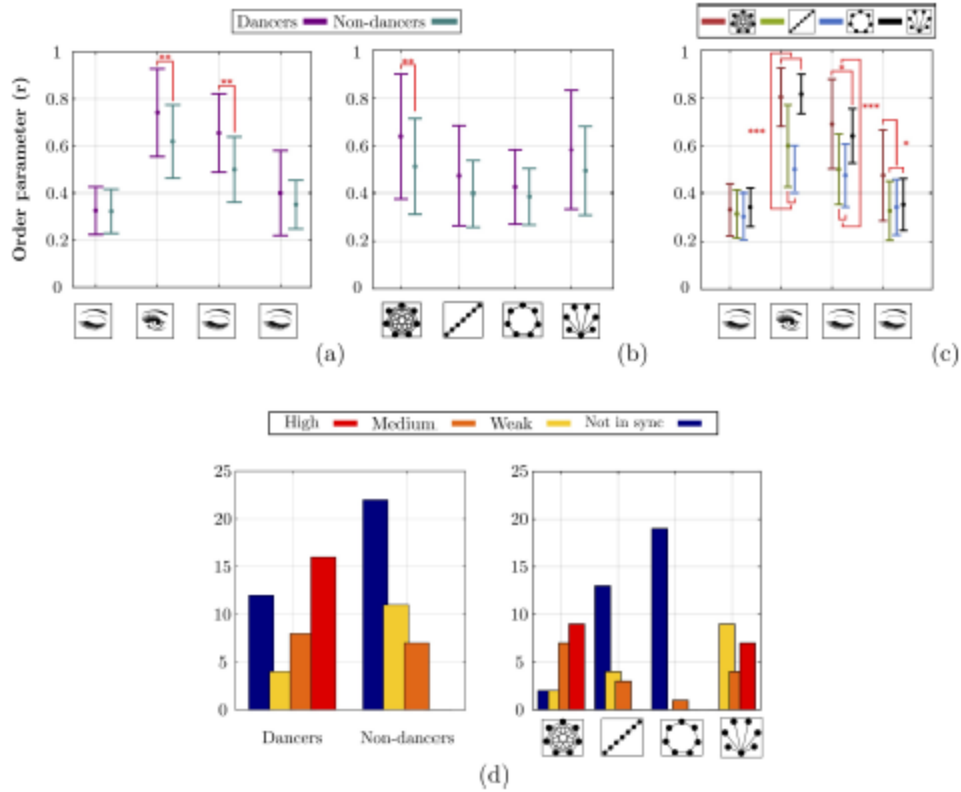


Figure 3: Main results of Experiment 2. Mean and standard deviation of Phase synchronization r in Experiment 2 as a function of (a) VisionExpertise, (b) ExpertiseTopology, (c) VisionTopology), $n = 320$; (d) distribution of phase synchronization levels across categories of robustness for Expertise (left panel) and for Topologies (right panel), $n = 80$.

We showed that our ability to move in unison is strongly influenced by our spatial configuration, similarity in behaviour, expertise and amount of visual exchange. In two experiments in which these factors, as well as their key interactions, were manipulated, we demonstrated that Complete and Star graphs were the most solid topologies prone to facilitating synchronized behaviours, reinforced by inertial homogeneity between participants and their expertise in perceptuo-motor synchronization. Importantly, we also demonstrated that group synchronization can be maintained for a certain amount of time (about 7 seconds) after informational exchanges have been interrupted, again more so in the two dominant topologies, and in a stronger way for experts. We investigated the origin of this effect by modelling our behavioural results with a simple ON-OFF dynamical model consisting in switching off the visual coupling and letting the individual dynamics relax to the initial oscillation frequency. This Static Coupling model was sufficient to partially capture our data. However, a memory effect had to be introduced in the model to account for the marked persistence of synchronization in eyes closed for two of the three homogeneity conditions, as well as for the coordination experts. Taken altogether, these results help to better understand why behavioural cohesion is easier to maintain when

perceptual exchanges are lost, more so in Path and Ring spatial configurations, and how perceptuo-motor expertise can reinforce this cohesion. How these multiple and co-existing configurations, within and across our senses, modulate our collaborative behaviours, in more naturalistic settings (richer perceptually and socially than pendulum operation in a laboratory), remains largely unknown and constitute a promising avenue for future research we aim to develop during EnTimeMent project.

References

- Damm, L., Varoqui, D., De Cock, V. C., Dalla Bella, S. & Bardy, B. Why do we move to the beat? a multi-scale approach, from physical principles to brain dynamics. *Neurosci. & Biobehav. Rev.* 112, 553–584 (2019).
- Kelso, J., Scholz, J. P. & Schöner, G. Nonequilibrium phase transitions in coordinated biological motion: critical fluctuations. *Phys. Lett. A* 118, 279–284 (1986).
- Nordham, C. A., Tognoli, E., Fuchs, A. & Kelso, J. S. How interpersonal coordination affects individual behavior (and vice versa): Experimental analysis and adaptive hkb model of social memory. *Ecol. Psychol.* 30, 224–249 (2018).
- Oullier, O., De Guzman, G. C., Jantzen, K. J., Lagarde, J. & Scott Kelso, J. Social coordination dynamics: Measuring human bonding. *Soc. Neurosci.* 3, 178–192 (2008).
- 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 (1990).
- Strogatz, S. *Sync: The emerging science of spontaneous order* (Penguin Group, New York NY, USA, 2004).

2.3.2.2. Phase II – Leadership patterns during synchronization in complex human networks.

For a full description please see: Calabrese, C., Lombardi, M., Bollt, E. et al. Spontaneous emergence of leadership patterns drives synchronization in complex human networks. *Sci Rep* 11, 18379 (2021). <https://doi.org/10.1038/s41598-021-97656-y>

Synchronization of human networks is fundamental in many aspects of human endeavour. Recently, much research effort has been spent on analysing how motor coordination emerges in human groups (from rocking chairs to violin players) and how it is affected by coupling structure and strength. Here we uncover the spontaneous emergence of leadership (based on physical signalling during group interaction) as a crucial factor steering the occurrence of synchronization in complex human networks where individuals perform a joint motor task. To date no systematic study was carried out to understand, in a controlled set-up, how influential each player is in steering the dynamics of a complex human network towards synchronization despite leadership emergence having been highlighted as a crucial phenomenon to understand such behaviour in

the animal world (e.g., Couzin et al., 2005, Stroeymeyt et al., 2011). In psychology and organization management, theoretical and empirical studies have focused on leadership emergence mostly from an abstract and language-based signaling perspective. Influence signaling is assumed to flow during social interactions by means of information exchanged through abstract language, whereby personal attributes may promote the election of a leader in the group (Winston & Patterson, 2006), and the diversity among individuals in cultural and social attitudes toward cooperation and fairness are crucial for the emergence of communities (Han et al., 2017). Here, we refer to a more primal, non-verbal and embodied form of interaction, in which leadership emerges out of motor coordination among the individuals in the group. The goal is to further improve the multi-disciplinary understanding of leadership as an inner natural trait characterizing all animal groups where communication does not necessarily rely on language (Couzin et al., 2002; Koorehdavoudt et al., 2016).

In two experiments (Phase I data) engaging participants in an arm movement synchronization task, in the physical world as well as in the digital world, we found that specific patterns of leadership emerged and increased synchronization performance. We report that leadership spontaneously emerges as an organizing phenomenon steering and enhancing group coordination. More strikingly, we unfolded three fundamental patterns through which leadership can appear in a group; a pattern where the leader, i.e., the most influential group member, is the person moving ahead of all the others (phase lead), pattern where the leader is the person moving behind all of them (phase lag), and one where leadership is shared between the two individuals bracketing group motion, i.e., leading ahead and lagging behind the other participants. We repeated the experiment via Chronos (Alderiso et al., 2017), a recently developed digital platform allowing players to participate remotely to the experiment via a computer terminal equipped with a leap motion controller (like in Phase I of this experimental work). We found that leadership emerges and enhances group coordination (see Figure 1 for summary of patterns identified).

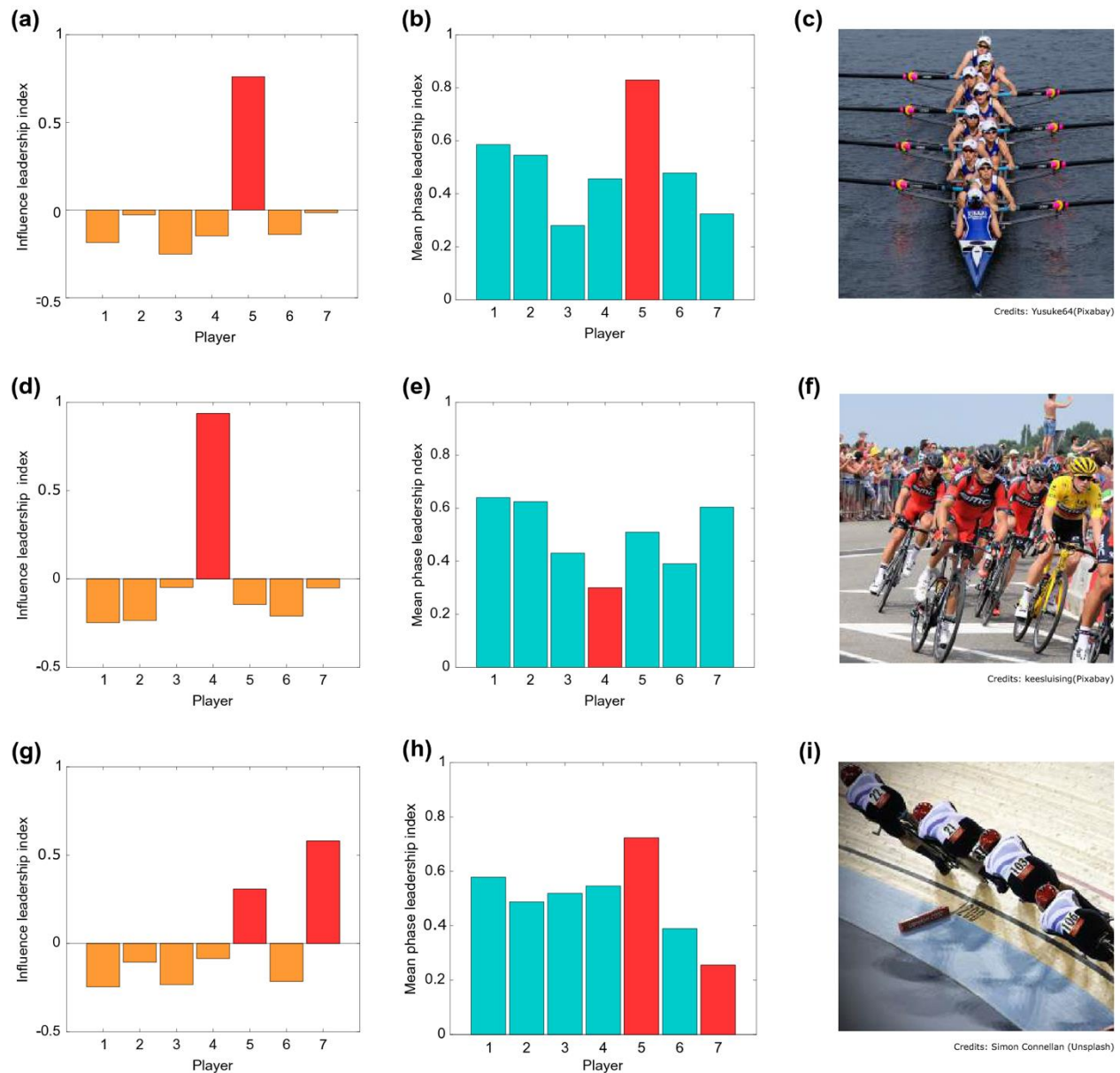


Figure 1: Description of the leadership scenarios. In the first column, each bar represents the value of the influence leadership index (see “[Methods](#)”). The central column shows the mean phase leadership index for each player. The red bars identify the relationship between these two metrics typical of each pattern (Pattern 1- top row, Pattern 2- middle row, Pattern 3- bottom row). The third column displays an example for each leadership scenario from sports world. In particular, (c) reports an example of the first scenario, where the role of a coxswain is to steer the boat and coordinate the rhythm of the team members. (f) Depicts professional road cyclists racing at the *Tour de France*. It is a paradigmatic example of leadership assigned to the participant lagging behind all the others, since the team leader typically drafts the wheels of her/his teammates (domestiques) but is pivotal for determining the team strategy. (i) An example of shared leadership observed during team pursuit in track cycling, where the fastest player in front sets the pace, but also accounts for the position of the third lagging player, since the final time for the team is taken when the third team member crosses the finish line.

Methods, experimental study and participants are identical to Phase I report (2.3.2.1), but here were re-analysed in terms of leadership emergence. For every individual, say i , we reconstructed from data the time-series of the phase of their motion. We then determined the phase leadership

index, say $H_i(t)$, of that individual in the group at every time instant. Namely, the value of $H_i(t)$ ranges from 1 to 7, with 1 (7) corresponding to player i indicates what group members are being followed (led) in phase by all the other player. We then computed the average phase index of individual i over the entire trial (H_i) to quantify their tendency to lead in phase. To complement the analysis and uncover who is influencing whom in terms of the information flow between group members, we evaluated next, from the time series of the phases of their motion, the causation entropy between every pair of agents in the group (Sun et al., 2015).

Then, we computed the net predicted information flow, NetCaus_i , from agent i to the rest of the group. This index defines the ability of each agent of influencing the motion of the whole group. The combined analysis of the mean phase ranking and of the net predicted information flow across all the groups evidences the three leadership emergence patterns described in this report (see full publication for the details how those indexes were computed along with the level of coordination). Data was analysed to extract the phases (or positions) of the players' motion and compute two metrics to assess their leader/follower roles. Namely, an influence index based on the computation of causation entropy, and a phase leadership index based on the phase analysis of the time series of the players' motion.

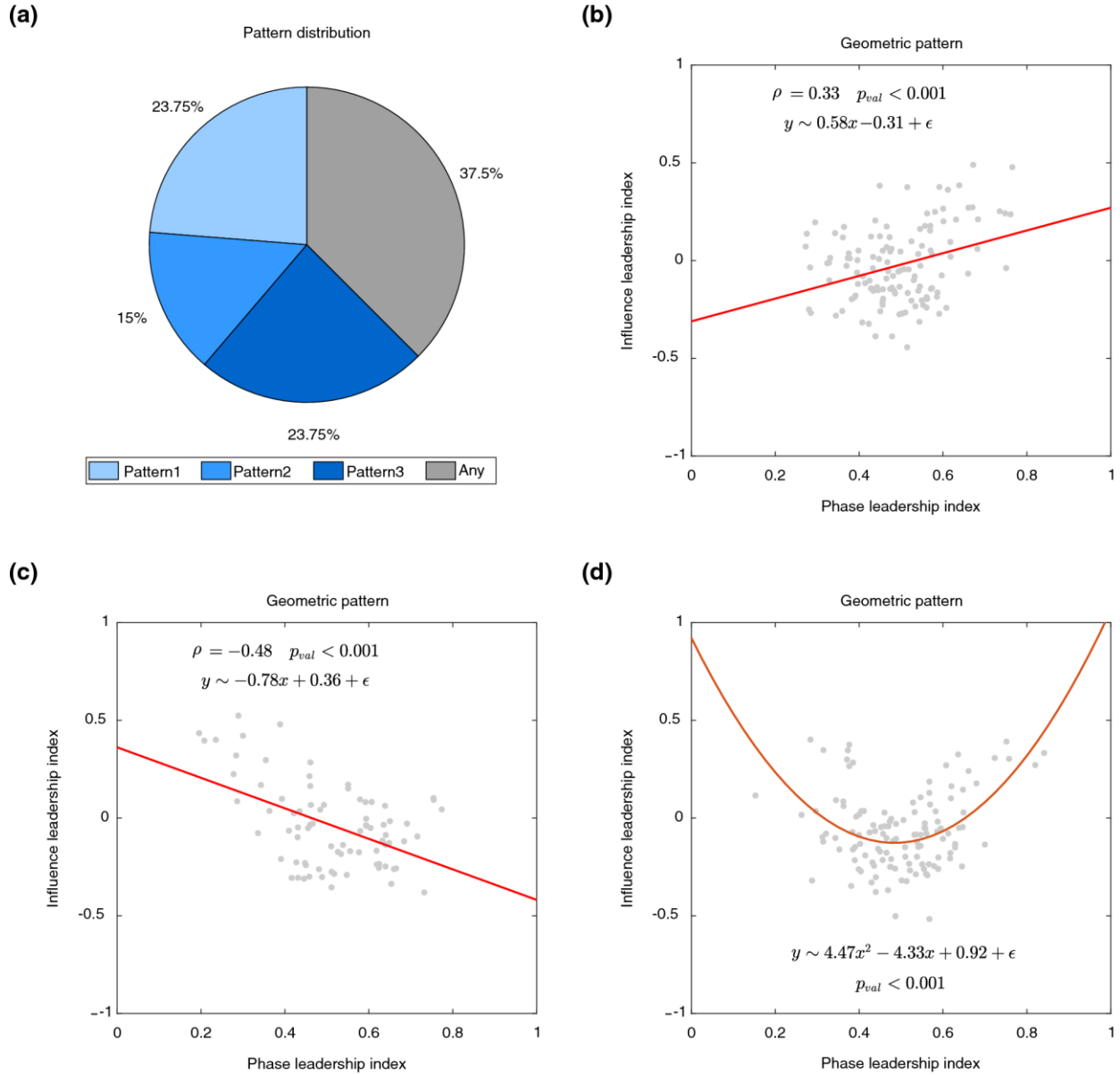


Figure 2: Distribution and characterization of the leadership scenarios. **(a)** The pie chart depicts how the three patterns were distributed among the trials of Experiment 1. A correlation analysis between the net information flow NetCaus and the mean position ranking $\square$ ----- in the trials where Pattern 1 and 2 were observed is reported in **(b)** and **(c)**, respectively. The linear fitting is represented by a red solid line. In Pattern 3, the relationship between NetCaus and $\square$ ----- is instead captured by a parabolic curve fitting, which is the red solid line in **(d)**.

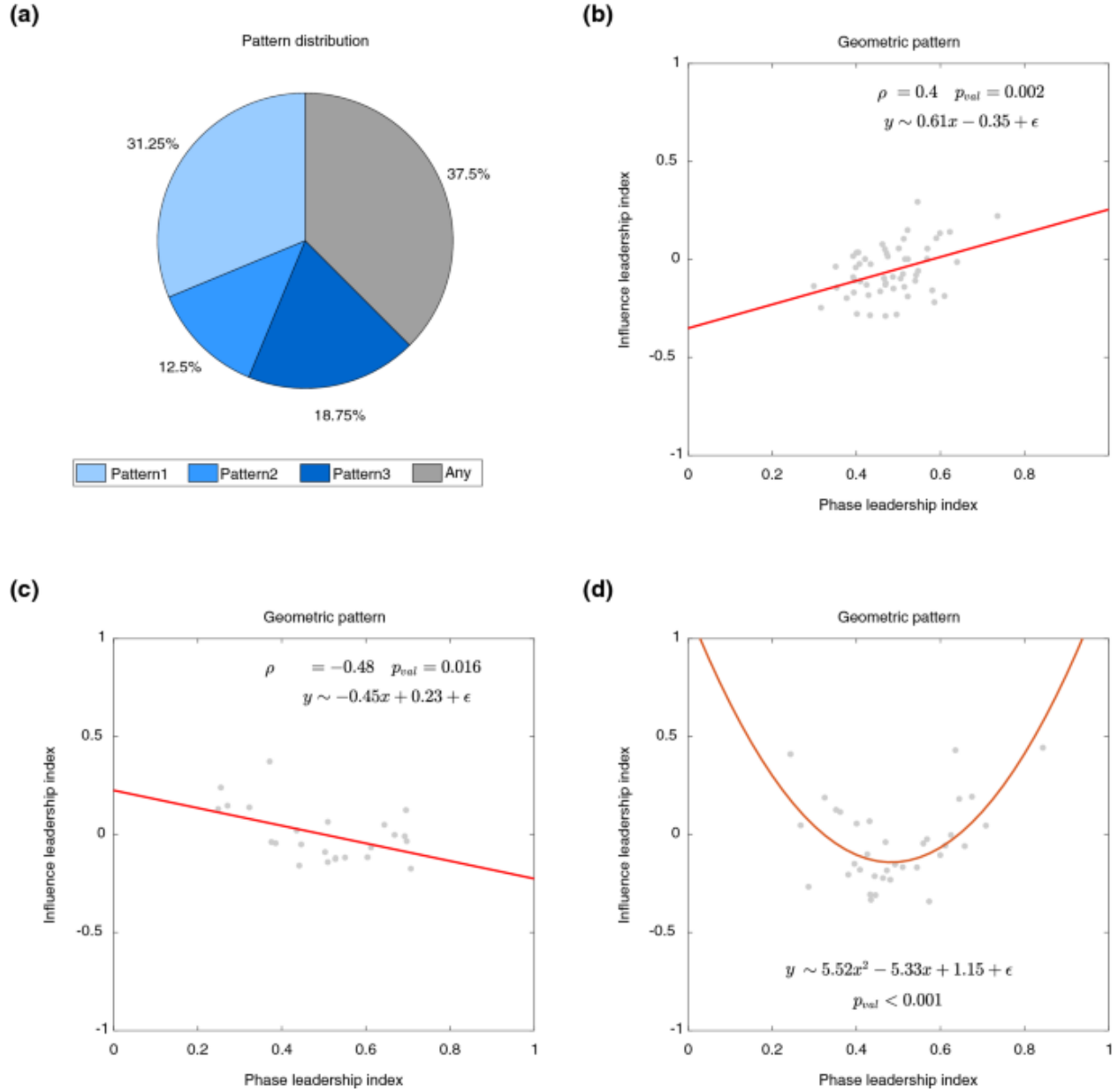


Figure 3: Distribution and characterization of the leadership scenarios in absence of social interaction. (a) The pie chart depicts how the three scenarios were distributed among the trials of Experiment 2. A correlation analysis between the net information flow NetCaus and the mean phase ranking $\square$ ----- in the trials where Pattern 1 and 2 were observed is reported in (b) and (c), respectively. The linear fitting is represented by a red solid line. In Pattern 3, the relationship between NetCaus and $\square$ ----- is instead captured by a parabolic curve fitting, which is the red solid line in (d).

As shown in Figure 2 and 3 we found that in over 60% of all trials, one out of three different leadership patterns clearly emerged.

In Pattern 1—“the fastest leads”—the individual who was the most influential, as assessed by computing the causation entropy between each pair of group members, was also the one leading the group motion in phase. In Pattern 2—“the slowest leads”—the most influential player was the one lagging behind all the others in phase. Finally, Pattern 3—“shared leadership”—was one

where the highest influence was shared by two players with the highest and lowest phase leadership index respectively.

The three leadership patterns we uncovered have analogues in configurations often found in nature and sports. For example, Pattern 1 is the one found in rowing competitions where rowers are driven and motivated by the coxswain, see Figure 1c. Pattern 2 resembles that used by teams in road cycling where partners work for the benefit of the team leader, who, although moving behind all the domestiques, determines the pace of the group, see Figure 1f. Pattern 3 is the same pattern emerging in track cycling where during team pursuit rounds the success of the group is influenced not only by the team member in the front position who sets the pace (*phase leader*) but also by the cyclist in the third position, who determines the final time by crossing the finish line, see Figure 1i. A geometric interpretation of these three patterns can be obtained by plotting the net predicted information flow against the mean ranking in phase for each of the three leadership scenarios, as shown in Figure 3.

A linear fitting shows that, in the trials where the phase leader emerges as the most influential (Pattern 1), ranking in phase and influence are positively correlated [$r(19)=0.33$, $R^2=0.11$, $p<0.001$]. Instead, when leadership is assumed by the player most lagging in phase (Pattern 2), the opposite is observed [$r(12)=-0.48$, $R^2=0.23$, $p<0.001$]. Interestingly, a linear fitting [$R^2=0.02$, $p=0.17$] cannot capture Pattern 3, where two players share leadership, and a parabolic interpolation is required to depict the coexistence of lagging and phase leaders [$R^2=0.24$, $p<0.001$].

Leadership emergence aids coordination

We found that one of the three leadership patterns described above emerged in a significant number of trials but not in all of them. Specifically, we detected one of the three scenarios emerging in 62.50% of all the trials, whereas in the remaining trials no consistent pattern emerged. We ascertained that the frequency of occurrence of one of the three patterns was different from chance by comparing it with the expected frequency of occurrence (28.57%) using a χ^2 -test [$\chi^2(3)=83.54$, $p<0.001$]. Most importantly, we found that the level of coordination in the group was significantly higher in those trials where leadership emerged than in those where it did not. In particular, we computed the group order parameter z (see Methods section in the publication for more details) to be equal to 0.73 on average when leadership emerged [$M1=0.73$, $SD1=0.17$] while being equal to 0.60 otherwise [$M2=0.60$, $SD2=0.17$]; such difference being statistically significant [independent t -test, $t(78)=3.37$, $p<0.001$, *Cohen's d*=0.78].

Group structure influences leadership emergence

We manipulated the group interaction structure to assess whether different topologies had an effect on leadership emergence. A χ^2 test revealed that the distribution of the three leadership patterns across the topologies is different from uniform [$\chi^2(3)=10.88$, $p=0.012$]. In particular, we observed a higher occurrence of leadership emergence (via one of the three patterns described in this paper) when all players were interacting with everyone else (complete graph topology,

Phase I, Figure 3c) or when a central player was interacting with the rest of the group (star graph structure, Figure 3f). These coupling configurations are also those that were found in previous work (Phase I, Alderiso et al., 2017) to facilitate the onset of motor coordination in the group [$\chi^2(3,76)=31.55$, $p<0.001$, $\eta^2=0.55$].

Leadership patterns persist in the absence of physical interaction

We repeated the experiment in the digital world without physical interaction, by means of Chronos, see Phase I for more information. We found again that, even in the absence of direct visual/auditory coupling among the players, one of the three leadership patterns emerged in 62.50% of the trials, see Figure 3a; a χ^2 -test confirming that the frequency of occurrence of the leadership scenarios was different from chance [$\chi^2(3)=18.81$, $p<0.001$].

Again, the average order parameter in the trials where these patterns are observed was higher than that observed in the remaining trials [$M1=0.68$, $SD1=0.16$ versus $M2=0.55$, $SD2=0.14$], and this difference was found to be significant [independent t -test, $t(30)=2.23$, $p=0.03$, *Cohen's* $d=0.82$]. Also, the geometric interpretation of the three scenarios in this second experiment was qualitatively the same as in the experiments where players were physically present in the same room.

We analyzed data from controlled experimental set-ups where a group of individuals was asked to perform an oscillatory task synchronizing its motion with that of others in the group. We found that leadership emerges spontaneously in complex human networks as an organizing mechanism facilitating synchronization, even when coupling among the players is digitally mediated. Three different patterns were detected through which leadership emerged. One, where the most influential player is the one ahead in phase, another, where the lagging one is the most influential, and the third, where a shared leadership arrangement emerges between the players furthest ahead or lagging in phase. Our results show that, as noticed for groups in the animal world, the emergence of leadership roles is a natural phenomenon occurring in human groups even when it is not solicited from external instructions or assignments. The *shared leadership* pattern that we observe in about 23% of all cases has been highlighted as relevant in neuroscience and psychology.

Our contribution perfectly matches the interest towards the expanding computational research in complex social network theory (Mahmoodi et al., 2020) and human interaction dynamics (Mao et al., 2017) that leverages the use of quantitative and computational methods to model, analyze, and interpret the mechanisms by which cooperation among humans evolves (West et al., 2016). Spontaneous collective behaviour emerges from individual internal decision making processes, each agent acting according to her/his own sensitivity to social opinion and group dynamics (Mahmoodi et al., 2017). Our unfolding of different leadership patterns represents a novel tool for to analyse individual strategies adopted in a group and to understand what motivates some players to behave as selfish agents or as resilient cooperators (Mao et al., 2017). In this study we have found three patterns of leadership, involving a subtle interaction between phase of the motion and amount of influence. Such patterns were independent of the presence or absence of

physical interaction, and persisted across manipulated spatial configurations. Our results shed light on the mechanisms that drive coordination and leadership in human groups, and are consequential for the design of interactions with artificial agents, avatars or robots, where social roles can be determinant for a successful interaction. Indeed, the leadership patterns that we highlighted might be used to program the motion of one or more artificial agents to enhance the performance of the human-robot group. This could be relevant, for instance, to aid the design of innovative rehabilitation procedures in those impairments where the imitation and movement synchronization mechanism fail (i.e., Virta et al., 2008).

References

- Alderisio, F., Lombardi, M., Fiore, G. & di Bernardo, M. A novel computer-based set-up to study movement coordination in human ensembles. *Front. Psychol.* **8**, 967 (2017).
- Couzin, I. D., Krause, J., James, R., Ruxton, G. D. & Franks, N. R. Collective memory and spatial sorting in animal groups. *J. Theor. Biol.* **218**, 1–11 (2002).
- Couzin, I. D., Krause, J., Franks, N. R. & Levin, S. A. Effective leadership and decision-making in animal groups on the move. *Nature* **433**, 513 (2005).
- Koorehdavoudi, H. & Bogdan, P. A statistical physics characterization of the complex systems dynamics: quantifying complexity from spatio-temporal interactions. *Sci. Rep.* **6**, 1–13 (2016).
- Han, X. *et al.* Emergence of communities and diversity in social networks. *Proc. Natl. Acad. Sci.* **114**, 2887–2891 (2017).
- Mahmoodi, K., West, B. J. & Gonzalez, C. Selfish algorithm and emergence of collective intelligence. *J. Complex Netw.* **8**, cnaa019 (2020).
- Mahmoodi, K., West, B. J. & Grigolini, P. Self-organizing complex networks: individual versus global rules. *Front. Physiol.* **8**, 478 (2017).
- Mao, A., Dworkin, L., Suri, S. & Watts, D. J. Resilient cooperators stabilize long-run cooperation in the finitely repeated prisoner's dilemma. *Nat. Commun.* **8**, 1–10 (2017).
- Stroeymeyt, N., Franks, N. R. & Giurfa, M. Knowledgeable individuals lead collective decisions in ants. *J. Exp. Biol.* **214**, 3046–3054 (2011).
- Sun, J., Taylor, D. & Bollt, E. M. Causal network inference by optimal causation entropy. *SIAM J. Appl. Dyn. Syst.* **14**, 73–106 (2015).
- Virta, M. *et al.* Adults with adhd benefit from cognitive-behaviorally oriented group rehabilitation: a study of 29 participants. *J. Atten. Disord.* **12**, 218–226 (2008).
- West, B. J., Turlak, M. & Grigolini, P. *Networks of Echoes* (Springer, 2016).
- Winston, B. E. & Patterson, K. An integrative definition of leadership. *Int. J. Leadersh. Stud.* **1**, 6–66 (2006).

2.3.4 Time to Sync

2.3.4.1 Phase I - EmoSync: human multi-scale group synchronisation

Interpersonal synchronization is the phenomenon that humans experience on a daily basis. While seated next to each other, and without realizing it, people tend to synchronize the rhythm of their leg swings (Schmidt, Carello, & Turvey, 1990), applause after a concert (Néda, Ravasz, Brechet, et al., 2000), rocking in chairs (Richardson, Marsh, Isenhower, et al., 2007), and side-by-side walking (Nessler & Gilliland, 2009). Studying humans during rhythmic interactions with others provides an ecological method for shedding light on social processes. Interestingly, movement in synchrony has a number of beneficial social consequences. For instance, interpersonal motor synchronization was found to increase social affiliation (Hove, & Rise, 2009+), to enhance self-esteem (Lumsden, Miles, & Macrae, 2014), and even to improve tolerance to pain (Tarr, Launay, Cohen, et al., 2015). Moreover, the historian McNeill (1997, p. 3) argued that the movement in synchrony—military drills in the everyday life of soldiers and the resulting emotional bond between the soldiers—have changed the course of human history.

The link between synchronization and the physiological arousal was empirically tested previously. Through shared perception humans can achieve synchronization (Oullier et al., 2008; Schmidt & O'Brien, 1997) and by watching a family member being engaged in a dangerous activity, the heartbeat of both tend to synchronize (Konvalinka et al. 2011); during a dialogue, the brain waves of the listeners are synchronized to the ones generated by the speaker (Hassan & Frith, 2016). What is not yet clear is how the emotional state impacts motor synchronization. Besides handful of studies (e.g., Fujiwara, & Daibo, 2018; Paxton, & Dale, 2013; Tschacher, Rees, & Ramseyer, 2014+; Varni, Volpe, & Camurri, 2010) the link remains largely unexplored and the results are inconsistent. While Fujiwara and Daibo (2018) did not find evidence for influence of affective valence on synchrony in a dyadic interaction, Varni and colleagues (2010) documented that comparatively to the negative emotion (i.e., anger), the positive emotion (i.e., pleasure) facilitated synchronization within a group of musicians. That is why we intended to systematically investigate how movement of a limb (e.g., Pollick et al., 2001; Sawada, Suda, & Ishii, 2003) in a positive and negative affective state alter group synchrony.

Thirty-nine right-handed participants ($M=25.44$, $SD=5.69$) that liked dancing participated in the study (21 females). Triads were randomly generated of people that were controlled for sex, level of expertise (Issartel et al., 2017) and acquaintance. When a triad was generated, an email was sent to the participants with names of other triad members and the participants needed to confirm no prior familiarity with others. However, if they were acquainted before, the triad was reassigned until no acquaintances were present in the group.

The task was developed on the basis of the mirror game (Noy et al. 2011), and its modification (Himberg et al., 2018) where the participants formed a triangle and performed an improvisation task with a limb. The experimental setup is illustrated in Figure 1.



Figure 1: The experimental setup in the Motion Capture room of Euromov laboratory.

No explicit instruction to synchronize was given to the participants. The instruction was: “to produce complex, varied and interesting movement with your right hand”.

There were three experimental conditions: Positive and Negative emotion induction as well as Neutral state which was also the control condition. The participants performed first 18 Solo trials and then 15 Group trials. The duration of each trial was 30 seconds. After each trial, participants were induced with an emotion. Right after the induction, the participants reported the psychological effect by filling a self-assessment manikin questionnaire (Bradley & Lang, 1994). To quantify the dynamics of the emotional bond created by synchronization, the dimensional model of emotion was chosen (Russel, 1980). The model consists of two scales: the arousal and the valence.

To make the emotion induction ecologically valid, the induction must be relevant to the participants. That is why prior to recruitment, the participants filled the questionnaire where they reported whether they liked dancing. The emotion induction consisted of manipulated feedback (Nummenmaa & Niemi, 2004). In essence, after each trial the participants were asked to rank each participant’s performance from the best to the worst. The participants performed ranking by using the Mentimeter application without being aware of the other’s voting. Then, they received one of three possible feedbacks: Positive - “NAME you were chose the best”, Negative - “NAME you were chosen the wors”, Neutral - “It’s a tie (first or last place) or second place”. Participants were evaluating subjectively who accordingly to their opinion performed the task the best. Participants evaluated last three performances in Solo trials and the last performance in Group. The feedback was manipulated and everybody was receiving the same feedback after each trial. The experiment was approved by the Euromov DHM Insitutional Review board #2002A.

In addition to psychological data, we recorded physiological data with the Delsys ECG Trigno sensor. The difference between two successive heartbeats—i.e., RR interval—was extracted with the sym4 wavelets that resembles the heart beats and therefore is suitable for R wave detection

of the heart signal. Finally, we registered behavioral data with the use of Vicon Motion Capture and 9 reflective markers placed on the participant to extract the cinematic data.

After the data were gathered, they were preprocessed. Below is the description and the visualization (Figure 2) of the order parameter extraction process. Firstly, the original velocities were extracted per each executed trial. Secondly, local detrending was performed to detect the oscillations with the use of inflexion points. Then, we analyzed the ondulatory part of signal to obtain oscillatory signal ranging from -1 to 1. Afterwards, we performed the Hilbert Transform to extract the phases. Finally, we quantified the Kuramoto order parameter on each trial of joint improvisation performance.

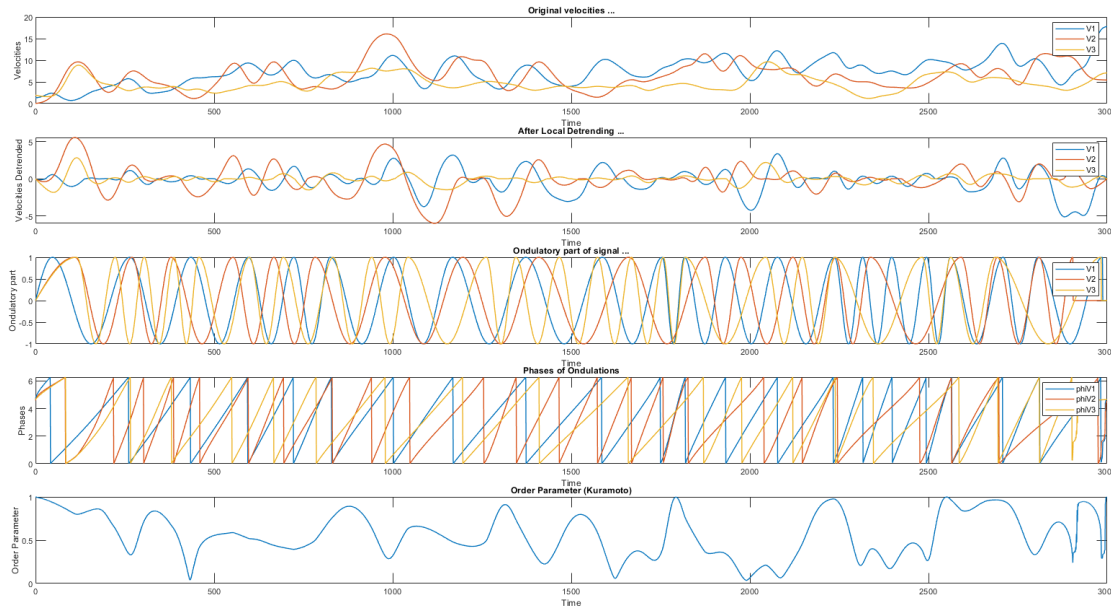


Figure 2: The Order Parameter extraction visualization.

We identified for all of the trials performed by a given Triad three levels of synchronization based on the IQR estimation. Accordingly, the order parameter of Q1 was selected as weak Sync, Q2 as medium Sync and Q3 as high Sync (Bardy et al., 2020).

After the data preprocessing, data were analysed. Hypothesis that Positive emotion would enhance synchronization was only detected in the behavioral data and there only the movement sync is reported here. We used the boxplot method and the interquartile ranges to check for outliers using the rstatistix R package. The method consists in finding the values bigger than $Q3 + 1.5 \cdot IQR$ and smaller than $Q1 - 1.5 \cdot IQR$ and identifying them as outliers. Then, the values bigger than $Q3 + 3 \cdot IQR$ and smaller than $Q1 - 3 \cdot IQR$ and identifying them as extreme outliers. We did not find any extreme outliers in our sample using the boxplot methods. Furthermore, the Time-In-Sync metric was normally distributed at each Emotion category, as assessed by Shapiro-Wilk's test ($p > 0.05$). Therefore, we were able to perform a parametric test, i.e. Repeated Measures ANOVA. The Time-In-Sync metric was statistically significantly different at the different Emotional Condition (Figure 1) during the trials, $F(2, 192) = 39$, $p < 0.001$, $\eta^2[g] = 0.29$.

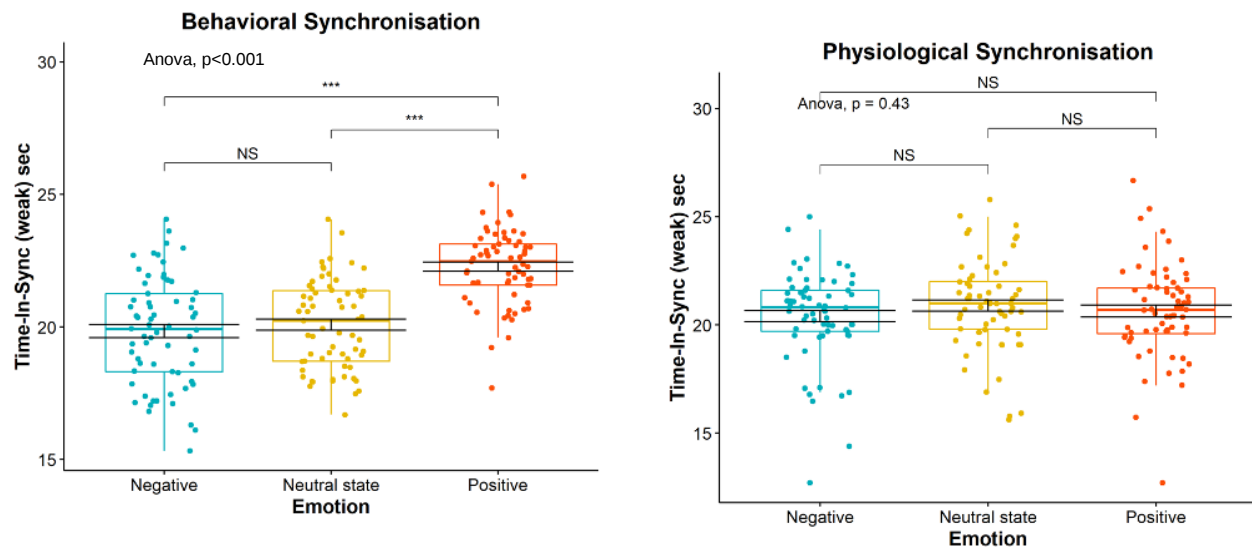


Figure 3: Boxplots with standard error bars displaying the distribution of the Time In Sync (TIS) metric of the behavioral (left) and physiological (right) dimensions of experiment. Synchronisation quantifies the correspondence of movement velocities (left) and heart rhythms (right) between three participants in a given triad. Each dot represents the result of TIS in one trial. For the behavioural aspect during the Positive Emotion trials, the participants spend more time in spontaneous synchronous interaction while Negative Emotion and Neutral state remained similar. For the physiological aspect, the experimental condition did not differ with respect to the time spent in synchrony.

References

- Bardy, B. G., Calabrese, C., Lellis, P. de, Bourgeaud, S., Colomer, C., Pla, S., & Di Bernardo, M. (2020). Moving in unison after perceptual interruption. *Scientific Reports*, 10(1).
- Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1), 49–59.
- Fujiwara, K., & Daibo, I. (2018). Affect as an antecedent of synchrony: A spectrum analysis with wavelet transform. *Quarterly Journal of Experimental Psychology*, 71(12), 2520–2530.
- Himberg, T., Laroche, J., Bigé, R., Buchkowski, M., & Bachrach, A. (2018). Coordinated Interpersonal Behaviour in Collective Dance Improvisation: The Aesthetics of Kinaesthetic Togetherness. *Behavioral Sciences (Basel, Switzerland)*, 8(2).
- Hove, M. J., & Risen, J. L. (2009). It's All in the Timing: Interpersonal Synchrony Increases Affiliation. *Social Cognition*, 27(6), 949–960.
- Issartel, J., Gueugnon, M., & Marin, L. (2017). Understanding the Impact of Expertise in Joint and Solo-Improvisation. *Frontiers in Psychology*, 8, 1078.

Konvalinka, I., Xygalatas, D., Bulbulia, J., Schjødt, U., Jegindø, E.-M., Wallot, S., . . . Roepstorff, A. (2011). Synchronized arousal between performers and related spectators in a fire-walking ritual. *Proceedings of the National Academy of Sciences of the United States of America*, 108(20), 8514–8519.

Lumsden, J., Miles, L. K., & Macrae, C. N. (2014). Sync or sink? Interpersonal synchrony impacts self-esteem. *Frontiers in Psychology*, 5, 1064.

McNeill, William H. 1995 *Keeping Together in Time: Dance and Drill in Human History*. Cambridge, MA: Harvard University Press.

Néda, Z., Ravasz, E., Brechet, Y., Vicsek, T., & Barabási, A. L. (2000). The sound of many hands clapping. *Nature*, 403(6772), 849–850.

Nessler, J. A., & Gilliland, S. J. (2009). Interpersonal synchronization during side by side treadmill walking is influenced by leg length differential and altered sensory feedback. *Human Movement Science*, 28(6), 772–785.

Noy, L., Dekel, E., & Alon, U. (2011) “The mirror game as a paradigm for studying the dynamics of two people improvising motion together,” *Proceedings of the National Academy of Sciences* (108) 52, 20947–20952.

Nummenmaa, L., & Niemi, P. (2004). Inducing affective states with success-failure manipulations: A meta-analysis. *Emotion* (Washington, D.C.), 4(2), 207–214.

Oullier, O., Guzman, G. C. de, Jantzen, K. J., Lagarde, J., & Kelso, J. A. S. (2008). Social coordination dynamics: Measuring human bonding. *Social Neuroscience*, 3(2), 178–192.

Paxton, A., & Dale, R. (2013). Argument disrupts interpersonal synchrony. *Quarterly Journal of Experimental Psychology* (2006), 66(11), 2092–2102.

Pollick, F. E., Paterson, H. M., Bruderlin, A., & Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition*, 82(2), B51-B61.

Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R. L., & Schmidt, R. C. (2007). Rocking together: Dynamics of intentional and unintentional interpersonal coordination. *Human Movement Science*, 26(6), 867–891.

Russell, J., A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39, 1161–1178.

Sawada, M., Suda, K., & Ishii, M. (2003). Expression of emotions in dance: Relation between arm movement characteristics and emotion. *Perceptual and Motor Skills*, 97(3 Pt 1), 697–708.

Schmidt, R. C., & O'Brien, B. (1997). Evaluating the Dynamics of Unintended Interpersonal Coordination. *Ecological Psychology*, 9(3), 189–206.

Schmidt, R. C., Carello, C., & Turvey, M. T. (1990). Phase transitions and critical fluctuations in the visual coordination of rhythmic movements between people. *Journal of Experimental Psychology: Human Perception and Performance*, 16(2), 227–247.

Tarr, B., Launay, J., Cohen, E., & Dunbar, R. (2015). Synchrony and exertion during dance independently raise pain threshold and encourage social bonding. *Biology Letters*, 11(10).

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

Varni, G., Volpe, G., & Camurri, A. (2010). A System for Real-Time Multimodal Analysis of Nonverbal Affective Social Interaction in User-Centric Media. *IEEE Transactions on Multimedia*, 12(6), 576–590.

2.3.3.2 ‘Spook and play’: What are the effects of emotional induction/perturbation using affective sounds from IADS-2 battery during playing a disembodied, multiplayer mirror game?

In EnTimeMent we embark on the scientific quest to use motion recordings as a primary source of information about one’s affective state through Machine Learning Algorithms (MLA) decomposing the signal and clustering it to match capabilities of detection pertinent to the human perceptual system. Neurologically healthy humans can detect emotions from other people and in the art (films and photographs, theatre, dance or music) through perception (seeing, hearing, feeling, smelling) and cognition (understanding). This capacity is rooted into the detection of signals embedded in static and dynamic patterns of emotional expressions and their comparison to others originating from their previous experience; innate knowledge plays a role as well (Matsumoto & Willingham, 2009). A plethora of research has provided different models of emotions (e.g., Ekman & Friesen, 1969; Izard, 1971; Plutchik, 1962; Russel & Fernandez-Dols, 1997) and have evidenced that agents can pick up emotions from facial expressions (e.g., Darwin, 1872, Ekman, Levenson, & Friesen, 1983; Cordaro et al., 2018) or from the way people walk (e.g., Roether et al., 2009) and reach for emotionally laden objects (Nogueira-Campos et al., 2016).

Emotional expression enriches communication with others and enables them to react in a timely manner to the changes in the environment. All animals are programmed by nature to move to obtain food (provide energy), avoid danger (survive) and reproduce (pass on genetic information) in cooperation or competition with others. From a biological perspective the primary function of emotion is to make the interactions with others more efficient, avoid threats and thrive in the surroundings (Barrett et al., 2019, Newman-Norlund et al., 2007). In the same vein, emotions can be understood as modes of action readiness relevant to the context a person is in (Jackendoff, 2007). The outcome emotion is derivative to the valence of the interaction (positive, negative) and level of arousal (low or high) (Kuppens et al., 2013, Russel, 1980) modulating the agent’s drive to approach or withdraw from source of emotion (object or person).

In a seminal study by Noy et al., (2011) mirror – game set-up was introduced to investigate how people are able to coordinate their movement in a novel and complex motor task (joint improvisation task) without a designated leader. This paradigm has been used by other studies successfully applying it as a canvas for studying interpersonal movement dynamics using an

index finger movement (i.e. Słowiński et al., 2016, Alderiso et al., 2017). Participants were asked to imitate each other's arm motion in such a way that they create one, harmonious and coordinated 'dance' like relationship between each other. This task was inspired by the professional practice that is commonly used by professionals for a warm up exercise in a theatre and during a dance practice.

In this study we propose to use an adapted, digital version of a multiplayer mirror game developed during the EuroMov ALTEREGO project - Chronos. It is an open-source software that allows egalitarian inclusion of participants regardless of their gender, age (social cues). This is achieved by movement of all participants being translated in real time from their personal space to gaming display (see Fig.1 for illustration). Their movement is usually captured by a [LEAP sensor](#) or trackpad/computer mouse (Alderiso et al., 2017). In addition visual contact is removed from participants (participants are separated by a visual barrier), and white noise delivered by the individual headphones to remove chance of auditory feedback. This way Chronos allows to test different synchronisation couplings, typologies (non-directional and Leader-Follower relationships) and interaction with virtual players. In addition, we aim to utilise a well established battery of emotion evoking sounds - The International Affective Digitized Sounds (IADS-2) (Bradley & Lang, 2007) as method of emotional induction for all participants delivered by personal headphones.

Our approach to this complex undertaking is built on the several assumptions, derived from the existing body of research suggesting that i.) movement qualities (i.e. speed, amplitude) are directly steered by emotional arousal; ii.) movement is a potent messenger of one's emotional state for the environment for others ; iii.) emotional qualities of the movement will be mirrored by others in a joint improvisation task iv.) different timescales might be employed for different emotional qualities. As a first step to set the background for the processing of the data, we need to define and detect a neutral state which corresponds from the informational point of view to a maximum entropy state. This means that when a person is in its neutral state, it is impossible to detect any emotional quality from the motion data because he or she does not experience any emotion providing a baseline comparison (Vergotte et al., 2018). In this scenario we propose that the neutral state will be derived from a trial accompanied by a sound of medium valence score and low arousal.

We tried to evoke positive and negative emotion using IADS-2 display (prior or during the launch of the movement trial). A plethora of research suggests that emotional content is best conveyed by the distal parts of the body (with relatively low energy expense cost for moving) during gross motor acts with high number of degrees of freedom (Roether, et al., 2009).

Before the start of the experiment each participant was seated in front of the computer display, separated from other players by physical distance and no direct visual contact (see Fig.1 as an example of set up). The workings of the LEAP motion sensor will be explained and demonstrated by the researcher - the required finger movement with right index finger and the optimal capture position. Participants will be instructed during the Solo trial to create an interesting non-periodic motion representing their motor signature and natural frequency ([Słowiński et al., 2016](#)) with

neutral sound being played. Then participants were asked to fill in two short questionnaires - selected items of PANAS (to assess their current emotional state prior to the experiment - Gaudrea et al., 2006 (French validation), and The Emotional Contagion scale (Doherty 1997, adapted by Alexander, 2017) to assess their baseline cognitive and affective empathy - Gilet et al., 2013 (French validation).

Procedure: After Solo trials participants were asked to synchronize their motion with that of the circles shown on their respective computer screen, representing the movements of the other agents topologically connected with them (Alderiso et al., 2017). However, players will have no global information of the topology of their interactions with other players. After each trial participants will be asked to fill in a digital assessment of the trial displayed with Mentimeter software online (using their personal phones) of their perception of the stimuli displayed prior/during the trial (to be decided post pilot of the experiment). We will use the three dimensions of pleasure, arousal, and dominance, the Self-Assessment Manikin (SAM), an affective rating system devised by Lang (1980). In this system, a graphic figure depicts values along each of the 3 dimensions on a continuous scale indicating emotional reaction to a stimulus presented. Although IADS-2 have been validated in multiple studies, we aim to control an emotional reply to stimuli that can be subjective - for example for sounds that aim to be relatively neutral (low arousal, medium valence) might evoke a reaction due to personal association linked to acoustic memory. We have obtained permission to use SAM scale (Fig. 2) for the EnTimeMent project from CSEA media (copyright owner).

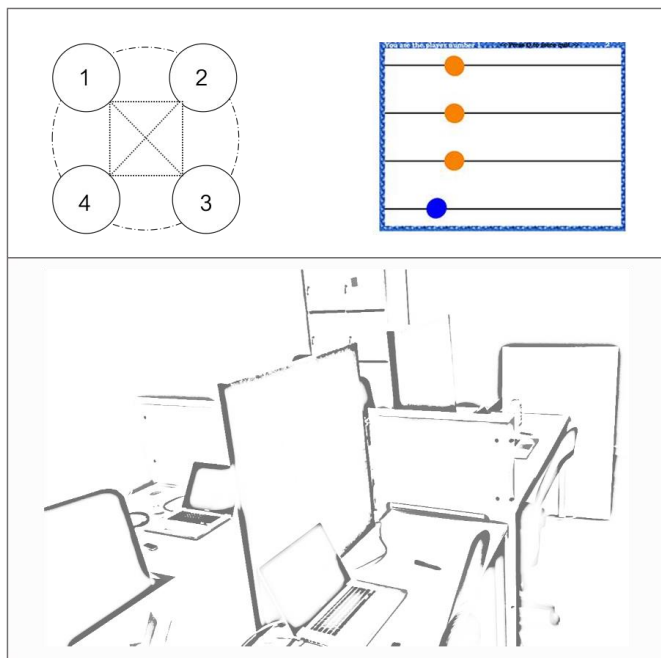


Figure 1: Participants performed 54 trials together, of 45 second duration each (this creates all possible combinations of the neutral, positive and neutral stimuli across the group (4participants x 9 possible alignments per condition x 3 types of stimuli). Top panel shows the complete (Bardy et al., 2020) coupling structure between participants using the digital display of avatars (blue is avatar of the participant – always position on the bottom of the screen, orange balls

depict avatars of other members of the group). Bottom panel depicts the experimental room with the shields between the participants to prevent view and real-time identification of the other members of the group.

Participants: 60 right-handed participants (28 females) were recruited from the student and staff pool of STAPS University of Montpellier and randomly assigned to 15 groups of 4 people mixed sex and age participants (age $M=26\pm 6$ years). Since the vision the Chronos display movement of other players will be anonymous, there is no need to control for the acquaintance level or participants being from the same age and sex group. All participants will be right handed with normal or corrected to normal vision and normal hearing.

Apparatus: The positional data was recorded with the LEAP motion sensors picking up the position of the right index finger of each participant as previously reported in Alderiso et al. (2017), with a frequency of 55Hz. The multiplayer mirror game setup will be running in a complete topology without a leader (meaning that all participants will be seeing each other's movement as illustrated on Figure 1). The acoustic emotional stimulus was delivered via headphones in each trial (randomized order across participants) with a delay of 6s to 8s post launch of each trial (this was done to avoid habituation of the participants to the occurrence of sound and create and temporal expectation, see Figure 2).

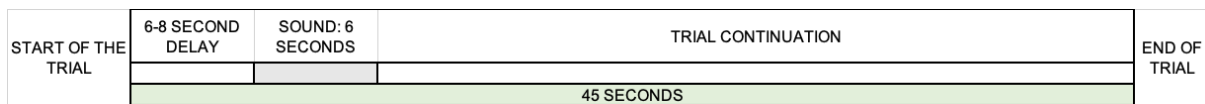


Figure 2: Delay in sound being played (each sound duration 6 seconds) was counterbalanced across trials – played between 6 and 8 seconds from onset of the trial to avoid participants habituation and confounding the perturbation created by sound in consecutive trials.

Independent variables: Three sets of emotional sounds were selected from the IADS-2 sample based on their ratings from Technical Report (positive -9 sounds, negative - 9 sounds and, medium valence/low arousal- 9 sounds (representative for neutral)). Each sound was used two times per participants to create required number of trials. Negative and positive sounds will differ in valence, but were matched for arousal using values from technical report with higher arousal scores being predesignated to negative sounds (IADS-2) as previously reported in the literature.

Analysis: Dependent variables: Mood and emotional contagion, handedness questionnaires and SAM responses were anonymised and matched with the same code that will be used in Chronos to record each participant's kinematic data. A Linear Mixed Model (LMM) was run to see whether gender or sensitivity to emotional contagion had an impact on the dependent variables computed from the data for each participant and group (synchronization measures were established using Kuramoto synchronization parameters and phase modelling as in experiments 2.3.2). Figure 3 depicts different behavioural observed across the groups in terms of the frequency of the movement and phase coordination (order parameter) (see Table 1 for overview of mean order parameter across the groups, where 1 is perfect synchronization).

Table 1

G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	G15
0,91	0,80	0,92	0,93	0,75	0,70	0,93	0,95	0,88	0,88	0,82	0,96	0,86	0,97	0,81

Note. Mean order parameter value per experimental group across all trial conditions (max=1, ideal synchronisation).

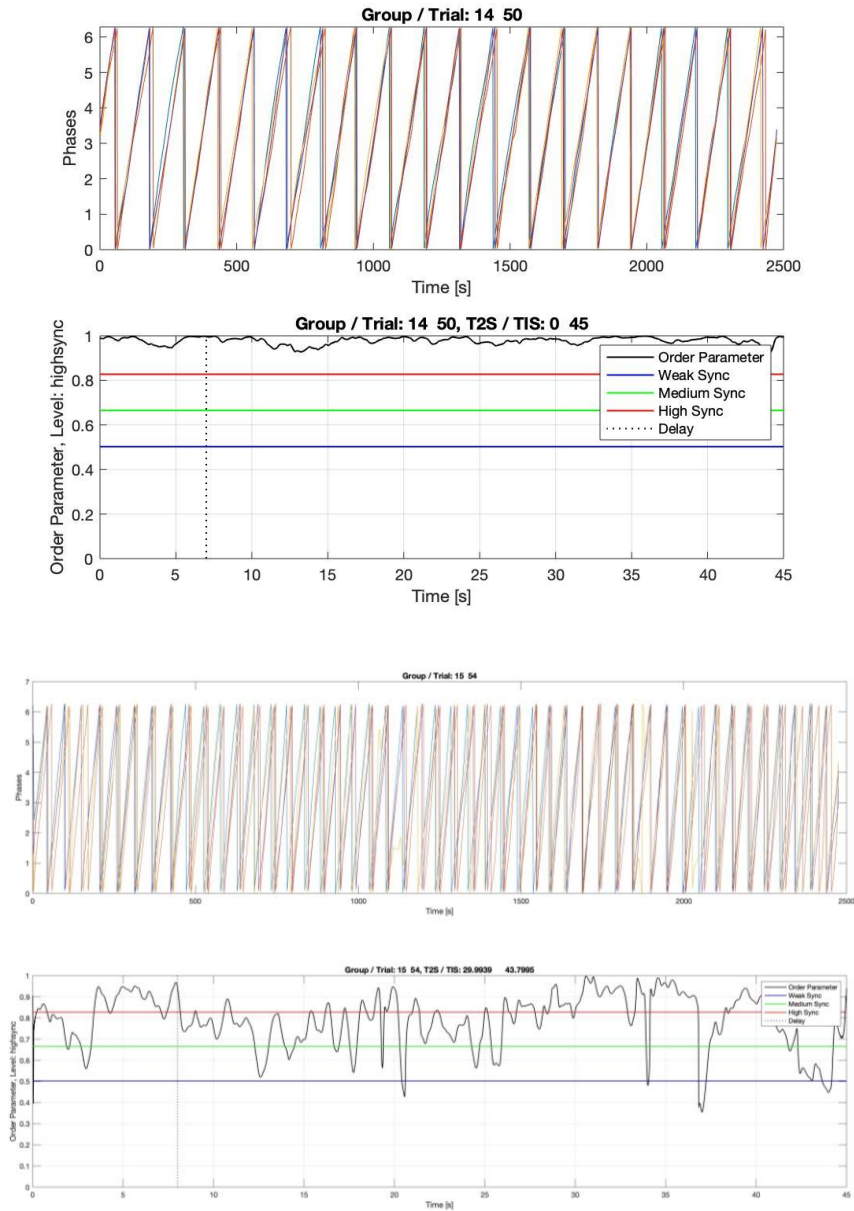


Figure 3: Levels of synchronization extracted from the experimental data (as quartiles of distribution of all order parameter values across all groups). Top panel depicts a group that moved relatively slower but maintained a high synchronization throughout the trial. In the bottom panel a group with more variable synchronization levels is presented, with synchronization dropping towards no synchronization in the middle of the trial and oscillating between medium and high order values.

In terms of the overall impact of emotion on the group and individual performance we are presenting here the main results, with models of best fit for the data. We have found no differences between females and males on their mood ratings prior to the experiment. The Emotional Contagion scale revealed that on average women scored higher than men, by independent t-test ($p=0.03$, Cohen's $d=0.6$), in particular on the scales of sensitivity to contagion of Fear ($p=0.005$, Cohen's $d=0.75$) and Sadness ($p=0.001$, Cohen's $d=0.9$), but not to Happiness, Love or Anger.

The spatiotemporal data was interpolated to 100Hz via spline interpolation. The order parameter achieved with the same processing pipeline as in 2.3.2 Dance to Sync, was further normalized with Fisher correlation z-scores. LMM analysis revealed that when three participants were induced with negative emotion as in opposition to none, the median order parameter drops (meaning synchronization is poorer), the time spent in the High synchronization decreases and standard deviation increases (group variability). The p-values estimated via t-tests using the Satterthwaite approximations to degrees of freedom. For the group analysis we dismissed the SAM ratings, as we concluded the averaging of scores per group is not representative as 'group emotion'. We have therefore used count of positive versus negative inductions per trial to model the random and fixed effects on order parameter. LMM model – Fisher transformation of order parameter {median per trial} post-sound induction period ~ Emotion (Negative count 3 [std $\beta = 0.16$, std CI=0.06-0.27, std $p = 0.003$], Marginal $R^2=0.02$, Conditional $R^2=0.61$); Time in High synchronization for the post-sound induction period {median per trial} ~ Emotion (Negative count 3 [std $\beta = 0.16$, std CI=0.28-2.48, std $p = 0.014$], Marginal $R^2=0.02$, Conditional $R^2=0.51$); *Standard deviation of the order parameter {median per trial}* ~ Emotion (Negative count 3 [std $\beta = -0.12$, std CI=0.01--0.01, std $p = 0.01$], Negative count 1 [std $\beta = -0.10$, std CI=-0.01--0.01, std $p = 0.01$], Marginal $R^2=0.02$, Conditional $R^2=0.57$).

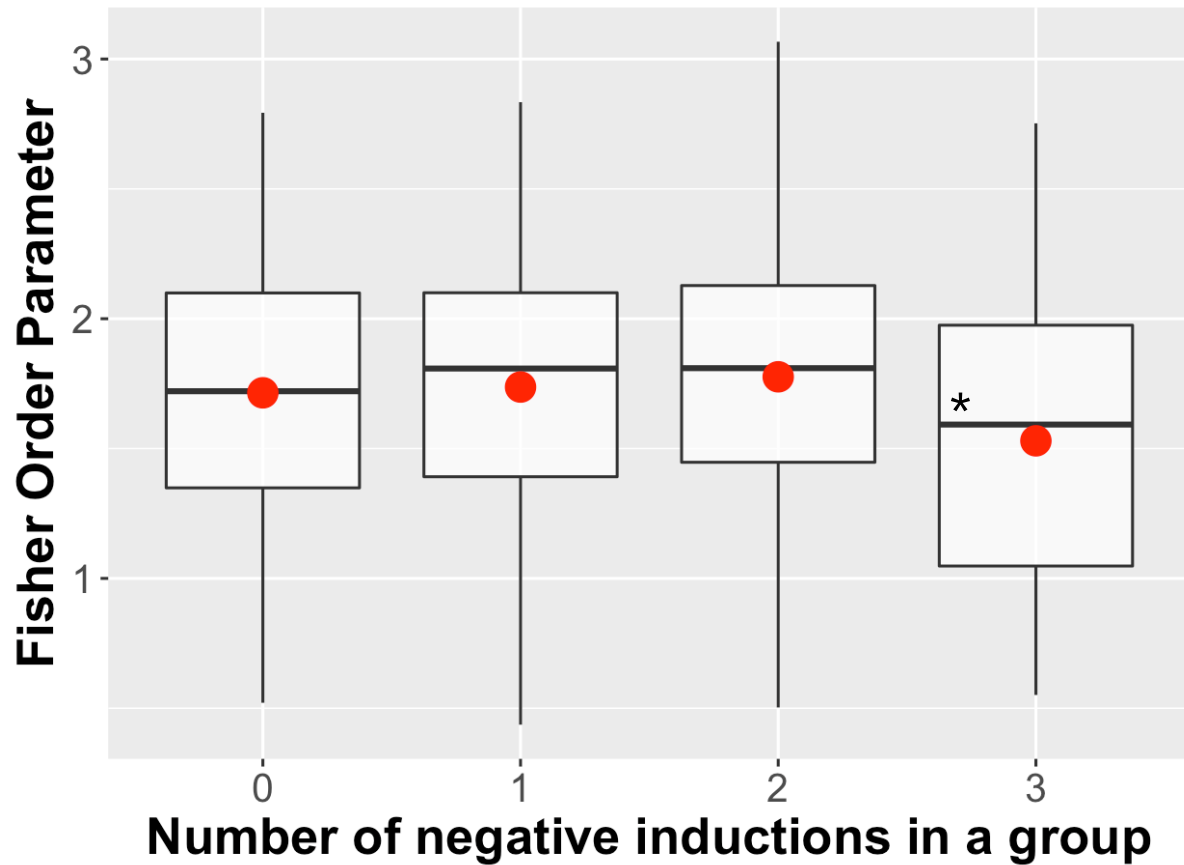


Figure 4: Mean order parameter per number of negative emotional inductions in the group.

Note:* Depicts the significant effect of 3 negative inductions in the group on the data.

In the Figure 5 we report the number of cycles per emotional condition, for the 35 second period post exposure to the sound. Participants were moving faster if more people were induced with positive emotion in the group, and slower with higher number of negative inductions in the group. LMM model – Number of Cycles per group trial ~ Emotion (Negative three/two [$\beta = 0.86/0.85$, $CI=0.53-1.20/0.62-1.08$, $p < 0.001$], Positive three/two [$\beta = -0.82/-0.56$, $CI=-1.16—0.48/-0.79—0.33$, $p < 0.001$], Marginal $R^2=0.003$, Conditional $R^2=0.88$).

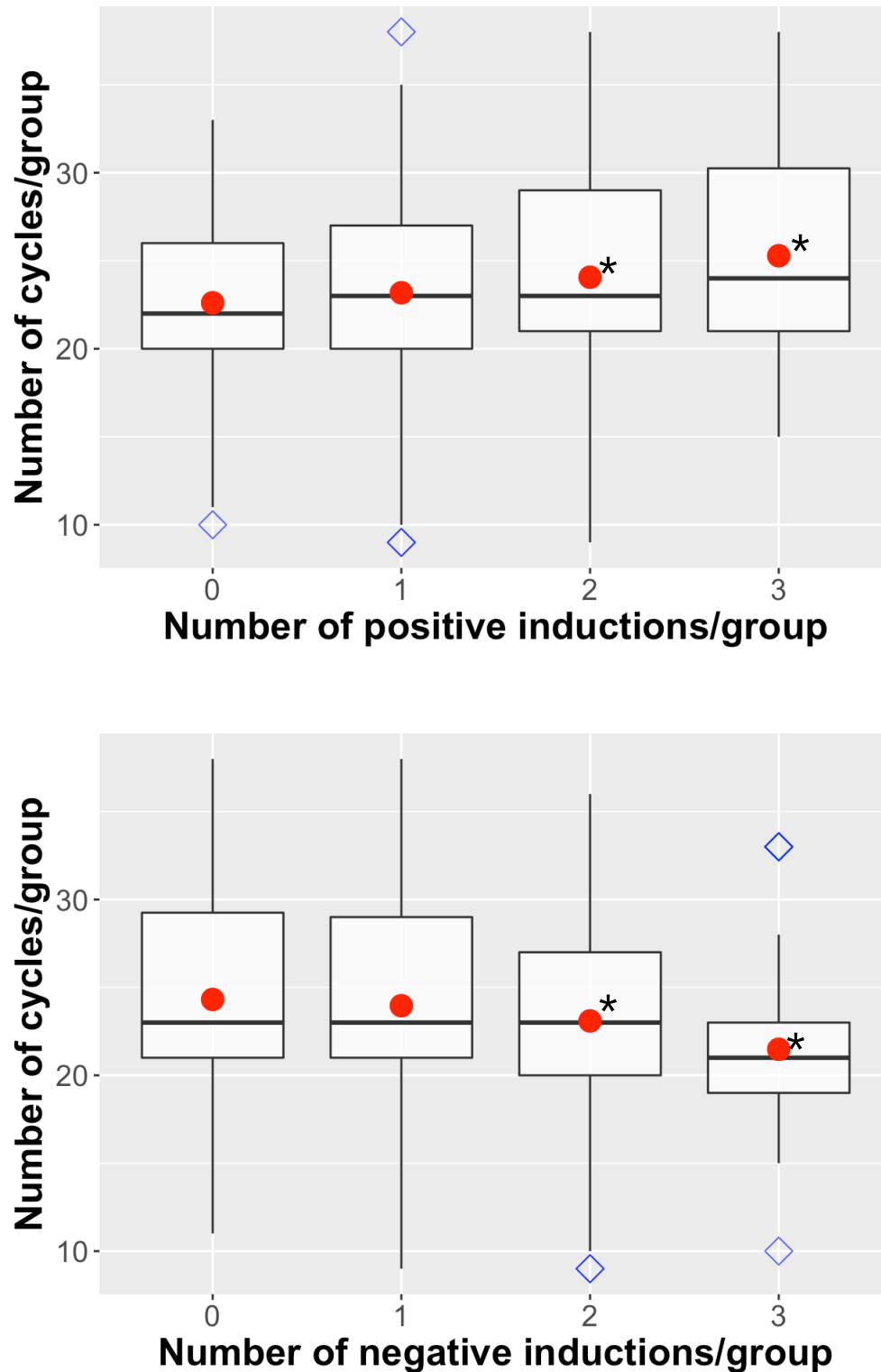


Figure 5: Number of cycles per positive inductions in the top panel per group, and negative bottom panel per group.

For the individual results, we have found that on average participant were moving faster induced with positive emotion and slower for negative induction. LMM model – Number of Cycles per trial

~ Emotion (Negative versus Neutral [std β = -0.10, std CI=-0.15--0.04, std p < 0.001], Positive versus Neutral [std β = 0.10, std CI=0.05-0.15, std p < 0.001], Marginal R^2 =0.006, Conditional R^2 =0.84).

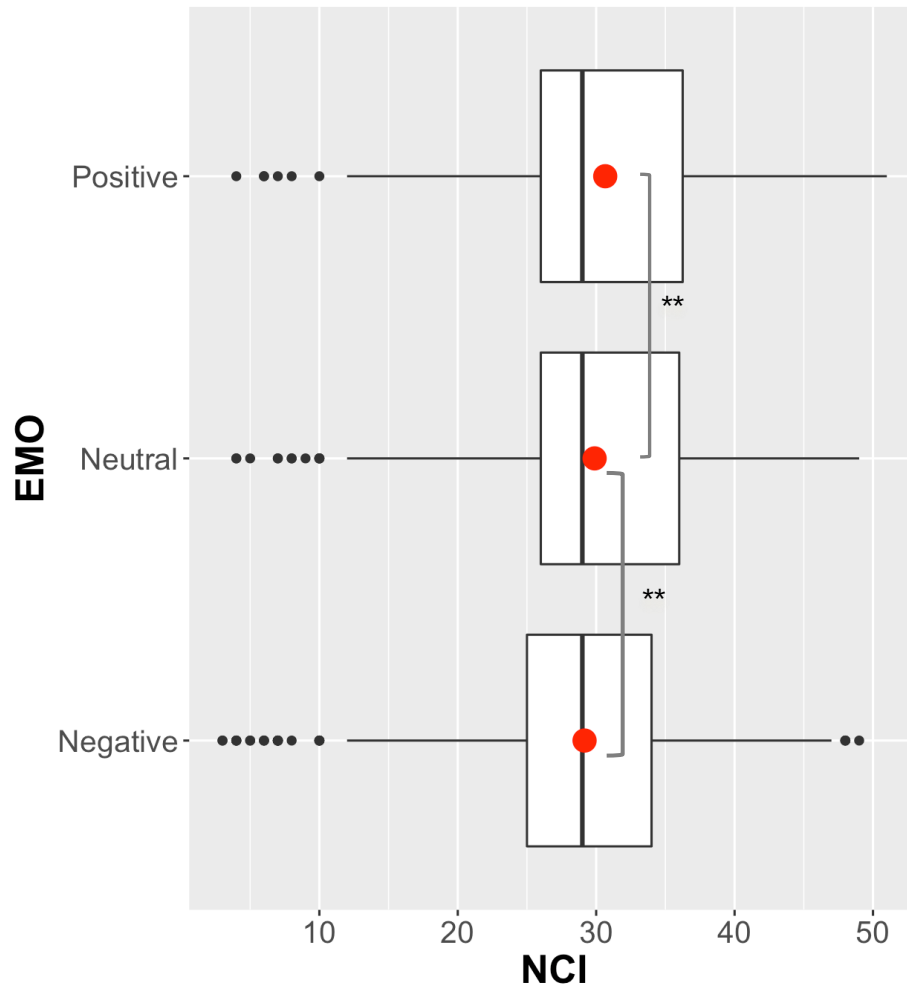


Figure 6: Average number of cycles per emotional induction condition. The number of cycles was significantly different for *neutral x positive* and *neutral x negative*.

Below we present synchronization measures – here index of synchronicity index calculated after Alderiso et al. (2017). Results showed that negative emotion was related to the lower value of individual synchronization index per LMM model – Synchronization Index ~ Emotion (Negative versus Neutral [std β = 0.06, std CI=0.02-0.09, std p = 0.001]) and women have not shown the same correlation between the scores on the Emotional Contagion scale, Synchronization Index ~ Emotion*Emotional Contagion score*Sex [std β = -0.05, std CI=-0.08--0.01, std p = 0.013, Marginal R^2 =0.05/Conditional R^2 =0.48] (see Figure 7).

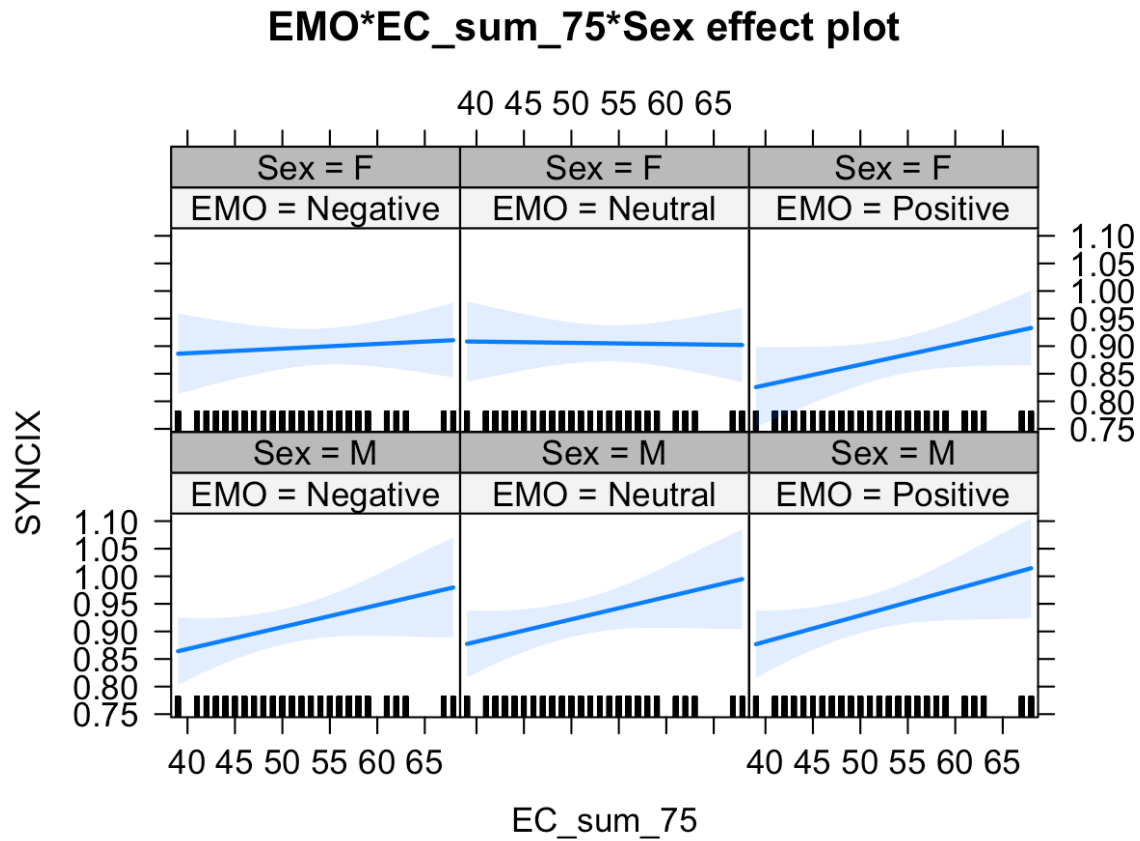


Figure 7: Individual synchronization index values for each emotional condition, per trial per participant with control for EC_sum_75, which denotes the sum of scores on the Emotion Contagion scale

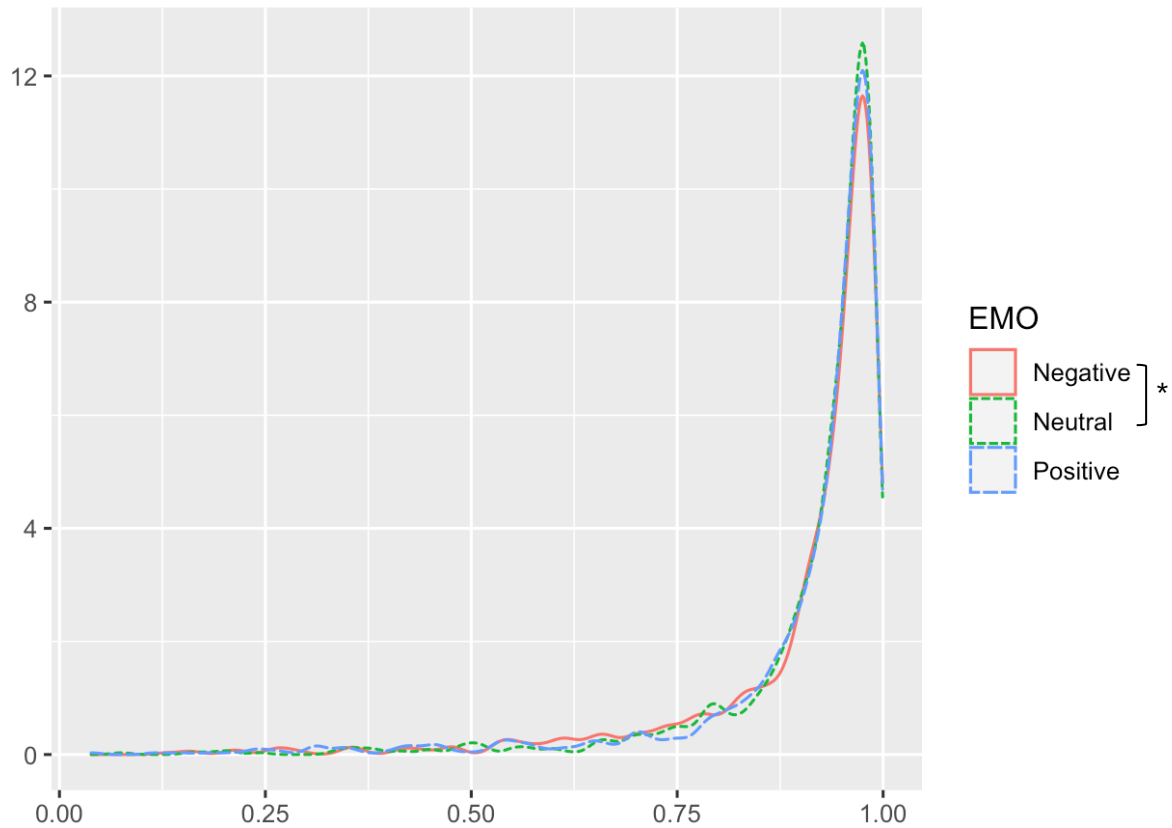


Figure 8: Distribution of the Individual Synchronization Indexes per emotional condition. The difference between neutral and negative was statistically significant, showing lower occurrence of higher synchronization index for negative induction trials.

Summary of results: From a theoretical standpoint, this study aims to help understand how positively or negatively loaded emotional stimuli impacts movement execution of a person in a group synchronisation context and whether it impacts the group as a social consequence of interaction. We found that negative induction in participants caused disruption in synchronization, measured by lower order parameter, less time spent in high synchronization levels and higher standard deviation. Similarly, both on the group and individual level negative induction was associated with slower movements and a smaller number of cycles of movement completed per trial, and positive one with faster movements and higher number of cycles. We have identified links of the scores of Emotional Contagion and synchronization index, with men maintaining the relationship for negative induction – higher scores align with better synchronization, whereas women not showing this tendency apart from positive induction only. This study paves a novel avenue for exploration how emotional context changes the way people perform a joint action task together.

References :

- Bardy, B.G., Calabrese, C., De Lellis, P. et al. (2020) Moving in unison after perceptual interruption. *Sci Rep* 10, 18032. <https://doi.org/10.1038/s41598-020-74914-z>.
- Barrett, L. F., Adolphs, R., Marsella, S., Martinez, M. & Pollak, S. D. (2019). "Emotional Expressions Reconsidered: Challenges to Inferring Emotion from Human Facial Movements." *Psychological Science in the Public Interest* 20 (1): 1–68.
- Cordaro, D. T., Sun, R., Keltner, D., Kamble, S., Huddar, N., & McNeil, G. (2018). Universals and cultural variations in 22 emotional expressions across five cultures. *Emotion*, 18, 75–93.
- Darwin, C. (1872). *The expression of emotions in man and animals*. London, England: Murray.
- Doherty, R. W. (1997). The emotional contagion scale: A measure of individual differences. *Journal of Nonverbal Behavior*, 21 (2), 131–154.
- Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: categories, origins, usage, and coding. *Semiotica* 1, 49–98.
- Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science*, 221(4616), 1208–1210.
- Jackendoff, R. (2007). *Language, consciousness, culture: Essays on mental structure*. Cambridge, MA: MIT Press.
- Gaudreau P, Sanchez X, Blondin J-P. Positive and negative affective states in a performance-related setting: *testing the factorial validity of the PANAS across two samples of French-Canadian participants*. *Euro J Psychol Assess* 2006; 22: 240– 249.
- Gilet, A.-L., Mella, N., Studer, J., Grün, D., & Labouvie-Vief, G. (2013). Assessing dispositional empathy in adults: A French validation of the Interpersonal Reactivity Index (IRI). *Canadian Journal of Behavioural Science / Revue canadienne des sciences du comportement*, 45(1), 42–48. <https://doi.org/10.1037/a0030425>
- Greco, A., Valenza, G., Citi, L. & Scilingo, E. P. (2017) Arousal and valence recognition of affective sounds based on electrodermal activity", *IEEE Sensors J.*, 17 (3), 716-725.
- Kuppens, P., Tuerlinckx, F., Russell, J. A., & Barrett, L. F. (2013). The relation between valence and arousal in subjective experience. *Psychological Bulletin*, 139(4), 917–940.
- Martin-Soelch, C., Stöcklin, M., Dammann, G., Opwis, K. and Seifritz, E. (2006) Anxiety trait modulates psychophysiological reactions but not habituation processes related to affective auditory stimuli", *Int. J. Psychophysiol.*, vol. 61, no. 2, pp. 87-97.
- Matsumoto, D., & Willingham, B. (2009). Spontaneous facial expressions of emotion of congenitally and noncongenitally blind individuals. *Journal of Personality and Social Psychology*, 96(1), 1–10.
- Nardelli, M., Valenza, G., Greco, A., Lanata, A. and Scilingo, E. P. (2015) Recognizing emotions induced by affective sounds through heart rate variability., *IEEE Transactions on Affective Computing*, vol. 6, no. 4, pp. 385-394, 2015.

Newman-Norlund, R., Noordzij, M., Meulenbroek, R., & Bekkering, H. (2007). Exploring the brain basis of joint action: Co-ordination of actions, goals and intentions. *Social Neuroscience*, 2(1), 48–65. <https://doi.org/10.1080/17470910701224623>

Noy, L., Dekel, E., and Alon, U. (2011). 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.

Nogueira-Campos, A. A., Saunier, G., Della-Maggiore, V., De Oliveira, L. A. S., Rodrigues, E. C., & Vargas, C. D. (2016). Observing grasping actions directed to emotion-laden objects: Effects upon corticospinal excitability. *Frontiers in Human Neuroscience*, 10(AUG2016). <https://doi.org/10.3389/fnhum.2016.00434>

Plutchik R. (1962) *The Emotions: Facts, Theories and a New Model*. Random House, New York.

Raymond, Alexandre (2017). La contagion émotionnelle : analyse de l'influence du niveau de stress perçu sur la sensibilité au processus de contagion émotionnelle. Mémoire. Rimouski, Québec, Université du Québec à Rimouski, Unités départementales des sciences de la gestion, p. 119.

Redondo, J., Fraga, I., Padrón, I. *et al.* Affective ratings of sound stimuli. *Behavior Research Methods* 40, 784–790 (2008). <https://doi.org/10.3758/BRM.40.3.784>

Roether, C., Omlor, L., Christensen, A., Giese, M.A. (2009). Critical Features for the Perception of Emotion from Gait. *Journal of Vision*, 8, 1-32.

Scherer, L. D., & Larsen, R. J. (2011). Cross-modal evaluative priming: Emotional sounds influence the processing of emotion words. *Emotion*, 11(1), 203–208. <https://doi.org/10.1037/a0022588>

Śłowiński, P., Zhai, C., Alderisio, F., Salesse, R., Gueugnon, M., Marin, L., ... Tsaneva-Atanasova, K. (2016). Dynamic similarity promotes interpersonal coordination in joint action. *Journal of the Royal Society Interface*, 13(116).

Soares, A.P., Pinheiro, A.P., Costa, A. *et al.* Affective auditory stimuli: Adaptation of the International Affective Digitized Sounds (IADS-2) for European Portuguese. *Behavioural Research* 45, 1168–1181 (2013). <https://doi.org/10.3758/s13428-012-0310-1>

Stevenson, R. A., & James, T. W. (2008). Affective auditory stimuli: Characterization of the International Affective Digitized Sounds (IADS) by discrete emotional categories. *Behavior Research Methods*, 40, 315–321. doi:[10.3758/BRM.40.1.315](https://doi.org/10.3758/BRM.40.1.315)

Vergotte, G., Perrey, S., Muthuraman, M., Janaqi, S., & Torre, K. (2018). Concurrent Changes of Brain Functional Connectivity and Motor Variability When Adapting to Task Constraints. *Frontiers in Physiology*, 9. <https://doi.org/10.3389/fphys.2018.00909>.

2.3.4 The sensitivity of intrapersonal coordination to interpersonal coupling

In D1.2 – previously entitled: *Similarity of motor signatures across multiple timescales in musical performers*

We adapt to complex and changing environments by finely coordinating multiple body parts at the same time. Doing so « in concert » with others enables very subtle forms of collaboration, such as playing together in a sport team or in a musical ensemble. However, coordinating the self and coordinating with others are most often investigated separately. In effect, motor coordination has long been studied in the sole context of individual actions, while a growing number of studies have recently focused on coordination at the interpersonal level (Schmidt & Richardson, 2008). From this perspective, intrapersonal coordination of movements can get « nested » within higher-order processes of interpersonal coordination (Ramenzoni et al., 2011). However, previous studies only involved postural or bodily symmetric tasks (ankles during gait, fingers during tapping, arms during precision tasks). They all did so in dyadic contexts, most often during visuomotor coupling tasks where goals predominantly targeted one level of coordination (either intra- or interpersonal) at the expense of the other. Yet, collaborative activities often involve different body parts (on top of postural demands) that move at distinct paces (unlike synchronized tapping or walking), and they can take place in larger groups of multi-modally coupled individuals whose performances are critical at both intra- and interpersonal levels. More complex and ecological tasks are thus required to grasp more realistically the sensitivity of intrapersonal coordination to interpersonal coupling.

A conducted musical ensemble is probably one of the best scenarios to tackle this issue (D'Ausilio et al., 2015; Volpe et al., 2016). Performers coordinate several body parts that move at multiple timescales to play complex musical patterns, and they aim at an exquisite temporal accuracy at both the individual and collective levels of coordination. Coupling between musicians is multimodal (auditory, but also visual, especially with a conductor), and body parts can serve different purposes, from instrumental gestures (those contributing to sound production) to so-called ancillary gestures (such as head movements, which communicate structure and convey expressivity).

To study it, we perturbed the network of sensorimotor communication of an orchestra: the visual coupling of first-section violinists with the conductor (normal condition) was substituted by a visual coupling with the second section of violinists (perturbed condition). Because the conductor plays a crucial role in organizing the temporal coordination of the ensemble (D'Ausilio et al., 2012), removing the vision of his gestures places a heavy burden on the timing of instrumental performance. We expected that, in order to focus on and refine the timing of their performance during perturbation, violinists would simplify the organization of their motor activity by increasing their intrapersonal coordination between ancillary (head) and instrumental (bow) movements.

A 17-piece orchestral ensemble, with 2 sections of violinists composed by 4 players each, and 2 different conductors were recruited for the experiment. The members of the orchestra were invited to perform in a concert hall (Casa Paganini, Genova) that was fully equipped with a motion capture system. The orchestra played a familiar excerpt from its repertoire - the opening of “Signor Bruschino” by Rossini -, which eschewed learning effects during the experiment. Participants were placed in two experimental conditions: a control condition (NORM) where all performers sat at their normal position in the orchestra, and a perturbed condition (PERT) where V1 turned their back on the conductor and faced V2 instead (see Figure 1).

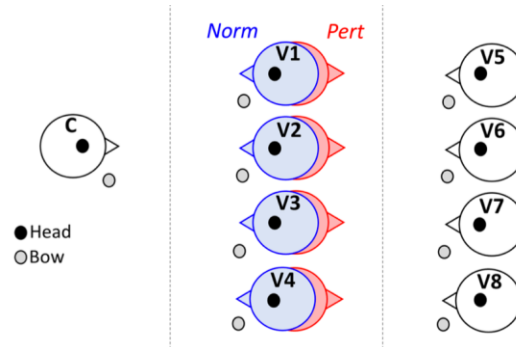


Figure 1: Position of the subjects in normal (NORM) and perturbed (PERT) conditions. Violinists v1 to v4 compose the first section of violinists (V1), and violinists v5 to v8 compose the second section (V2). The main change was introduced in the visual coupling of V1, shifting their visual field away from the conductor and toward the second section.

Movement data were collected with a Qualisys motion capture system equipped with 7 cameras and integrated with the EyesWeb XMI platform (http://www.infomus.org/eyesweb_ita.php). Violinists and conductors each wore a cap with three passive markers of the Qualisys motion capture system (positioned at Pz, F3 and F4 in the 10–20 electroencephalographic system), and another marker was placed on the tip of their bow and the conductors’ baton. Data tracking was done by the Qualysis Track Manager software, with a sampling rate of 100 Hz. An audio recording of the ensemble was also recorded and synchronized with the motion data.

We extracted the velocity time series of the head and the bow by computing the Euclidean distance between the successive positions of their associated markers (evenly space by 10 milliseconds). Data of each of the 4 selected passages of data were normalized to z-scores and filtered with a zero-phase second-order Butterworth bandpass filter between 0.5 and 12 Hertz. The cut-offs were chosen to reflect the main rhythmical values played with the bow, but also because the length of some of the passages were too short to capture lower frequency components.

Windowed cross-correlation analysis

To gauge V1’s overall intrapersonal coordination, we looked at the extent to which head and bow motion varied together. To do so, we performed windowed cross-correlations between their respective velocity time series. Cross-correlation coefficients were computed at up to 6 lags (± 60 ms).

To compare results across conditions, we extracted 5 indices from the cross-correlation functions obtained in each 1-s window, and we averaged those indices by trial for each performer. First, we extracted the correlation coefficient at lag-0. This measures the extent to which head and bow synchronously moved together. Second, since head and bow motion could also be coordinated with a (variable) delay, we extracted the highest coefficient value found among the 13 possible lags (-60 to +60 ms by steps of 10 ms) at which cross-correlations were computed. This estimates the magnitude of coordination between head and bow motion, irrespective of delay. We extracted these peak coefficient values in the surrogate trials as well and used the 95th percentile of the resulting distribution to establish a baseline above which lagged correlations are considered non-random. Importantly, Fisher z-transformation was applied to all coefficients values before performing any averaging and further statistical analysis. To present the results numerically and graphically (and when no more processing was to be performed), we used hyperbolic tangent transformation to revert the values back to the scale of correlation coefficients.

Next, we extracted the lags at which peak correlations were observed. This quantifies the magnitude of the delay at which time series are maximally correlated, as well as its direction (which time series precedes the variation of the other). To isolate the magnitude of this delay irrespective of its directionality, we then extracted the absolute lags at which peak correlations occurred (from 0 to 60 ms). Finally, we computed the standard deviation of the lags at which each peak correlations occurred. This quantifies the stability over time of the delay at which correlations were maximal, and, as such, the stability of the temporal coordination between head and bow time series.

Movement amplitude and spatial dispersion analysis

To better understand the factors that underlie a potential change in the intrapersonal coordination of head and bow, we gauged the amplitude of their displacements. To do so, we first measured, in each of the above windows of analysis, the spatial dispersion and the volume these displacements covered. Spatial dispersion was estimated by computing the mean (Euclidean) distance between all positional datapoints (i.e., all positions in space that head and bow respectively visited). We then computed the volume contained by the 3D convex hull of head and bow spatial trajectories (i.e., the volume of the smallest possible polyhedron that contained all data positions). This indicates the amount of space covered by head and bow motion trajectories. Both these indices were averaged for each trial of each performer.

Power spectral density analysis

To better understand the temporal structure of bow and head motions, we examined them separately in the frequency domain by computing their respective power spectral density (PSD, using the `pwelch` function in Matlab). Since the rhythmical content of the musical score is changing over time, and since tempo fluctuates over the course of the performance (with a notable shift toward acceleration in the last portion), we proceeded by short windows. This helped us focusing on the frequency range within which instrumental motion was prominent (between 1 and 8 Hz approximately, which roughly correspond to whole and eighth notes respectively). To capture the lowest component accurately enough, we used windows of 3 seconds (i.e., 300 data points). The

window of analysis did not overlap. Windows were then zero-padded to obtain time series of 512 points to be submitted to power spectral density analysis.

Power correlation analysis

To verify if changes in head motion frequency composition could reflect the mimicking of the rhythmical movements of the bow, we constructed time series of the power observed at each metrical timescale for both head and bow motion. To do so, we extracted the power observed at each relevant frequency bin in each window of PSD analysis performed above. The resulting time series thus illustrate the evolution of power at each relevant metrical timescale. Then, we computed correlations between head and bow time series, which indicates the extent to which the power of head and bow motion vary together at each metrical timescale.

To verify if changes in head motion frequency composition could be due to the visual coupling with V2 that occurred during PERT, we first computed PSD for V2's head and bow motion. Then, we repeated the above procedure and extracted the time series of the power of V2's head and bow motion at each metrical timescale. Finally, at each timescale, we computed the correlation between the time series of power observed in the head motion of each V1 performer and the time series of power respectively observed in the head and bow motion of the V2 performer each V1 performer was facing. This should indicate the extent to which V1's head motion and V2's head and bow motion evolved similarly, suggesting (or not) an informational coupling between the two sections.

Relative phase analysis

To evaluate how bow and head were coupled with respect to the multiple timescales at which their motion was organized, we examined their phase relationships in all musically relevant frequency bands (those corresponding to the metrical levels designated above). The centroid of each frequency band was derived from the ongoing local tempo computed above for each window of analysis. We then filtered the time series with a second-order Butterworth bandpass filter with zero-phase distortion, with cut-offs set at + and - 0.5 Hz away from the centroid of each relevant frequency band. The resulting head and bow time series were subjected to Hilbert transform, from which we derived their respective phase angle time series. On this basis, we then estimated instantaneous phase differences between them (or RP for relative phase). Finally, we extracted 3 indices from the RP time series at each timescales (averaged across windows for each trial of each performer): 1) the vector length (VL) which gauges the stability of the phase relationship (i.e., how few the RP varies over time) 2) the mean angle (MA), which quantifies the average phase difference between the time series 3) the mean absolute angle (MAA) which captures the average phase difference regardless of its directionality (i.e. which series precedes which).

Statistical analysis

Since criteria such as the normality of the distribution and the homoscedasticity of the data could generally not be assumed, and because of the repeated-measures design, we used non-parametric Friedman tests for the statistical analysis of all the above indices. Given the small population of subjects that is inherent both to the ecological design of the experiment and to the actual composition of a (chamber) orchestra, such a test offers the opportunity to remove the effects linked to the subjects' factor and therefore allows to address directly the difference

between our 2 conditions of interest (NORM and PERT). To perform this test, we collapsed the results obtained across the 2 different conductors. In effect, this factor did not introduce any difference in the main analyses we intended to perform (the overall intrapersonal coordination measured with windowed cross-correlations). Since a few significant differences have been observed for other variables, we report the comparisons between conductors for all variables and for each experimental condition in the supplementary materials. For the main analyses, collapsing data across conductors resulted in matrices of 4 blocks (corresponding to 4 violinists) that each contain 6 replicates (3 trials by 2 conductors) as row effect, and 2 experimental conditions (NORM vs PERT) as the column effect to be tested.

Summary of results:

Windowed cross-correlation analysis

To gauge V1's overall intrapersonal coordination, we performed windowed cross-correlations on head and bow velocity time series (average cross-correlation functions are presented in figure 2). Lag-0 correlation coefficients were significantly higher in PERT ($r = .23$, ± 0.07) than in NORM ($r = .12$, ± 0.05 ; $\chi^2 = 12.98$; $p = 0.0003$). Head and bow thus moved more synchronously during PERT. Peak coefficients were also significantly higher in PERT ($r = .42$, ± 0.07) than in NORM ($r = .36$, ± 0.05 ; $\chi^2 = 12.41$; $p = 0.0004$). Head and bow thus tended to move more similarly during PERT, irrespectively of the lag difference between the 2 time series. On average, peak correlations were observed near lag-0 (NORM: -3.19 ± 9.57 ms; PERT: -7.08 ± 8.29 ms), indicating that performers tended to synchronize their head and bow movements. Lags associated to peak correlations did not significantly differ across conditions ($\chi^2 = 0.41$, $p = 0.52$). However, there was a marginal tendency for absolute lags to be smaller in PERT (32.92 ± 9.90 ms) than in NORM (38.89 ± 8.33 ms; $\chi = 3.41$; $p = 0.0647$). Thus, head and bow tended to move more synchronously during PERT than NORM. Finally, the lag at which peak correlations occurred was less variable in PERT (33.63 ± 11.62 ms) than in NORM (42.09 ± 8.65 ms; $\chi^2 = 5.77$; $p = 0.0163$). The temporal coordination between head and bow motion was thus more stable (i.e. less variable) during PERT. In sum, head and bow overall intrapersonal coordination was stronger (higher correlation coefficients), more in phase (closer to lag-0) and more stable (less variability in the lags of peak correlations) during PERT.

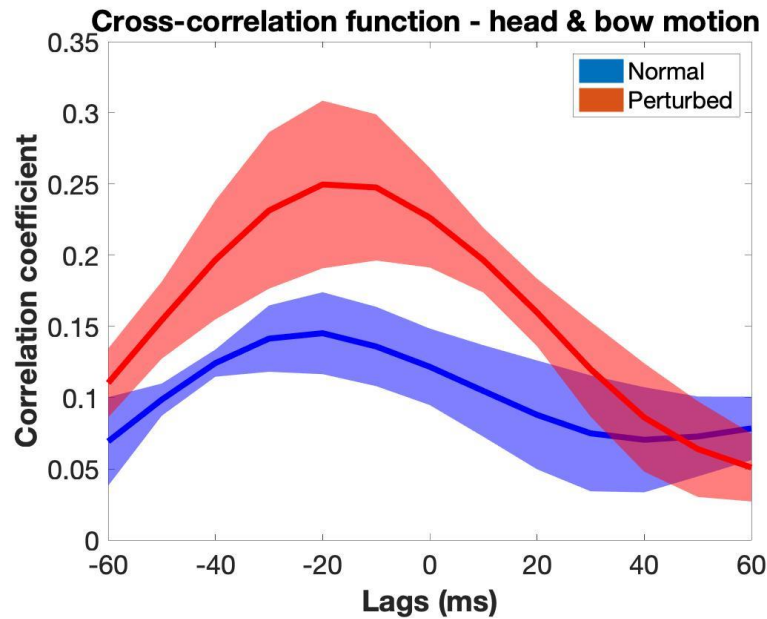


Figure 2: Average cross-correlation functions between head and bow motion in the NORM and PERT conditions (shaded areas represent standard error). Correlation coefficients were higher in the PERT condition than in the NORM condition, suggesting that head and bow were more strongly coupled during PERT.

Movement amplitude and spatial dispersion analysis

To investigate the kinematic changes that underlain differences in intrapersonal coordination, we gauged movements amplitude by measuring their spatial dispersion (mean inter-distance between all datapoint positions) and the volume covered by head and bow motion (convex hull). Mean inter-distances did not differ across conditions for the bow ($\chi^2 = 2.91$; $p = 1$) but they were significantly larger in PERT than in NORM for the head ($\chi^2 = 23.08$; $p = 0.000002$; see table 1 and figure 3). Head (but not bow) motion trajectory thus visited positions that were more spread in space during PERT. The volume contained in the convex hull of bow trajectories did not differ between conditions ($\chi^2 = 0.06$; $p = 0.81$), but it was significantly larger in PERT than in NORM for head trajectories ($\chi^2 = 25.44$; $p = 0.0000005$; see table 1 and figure 3). Motion of the head, but not of the bow, covered larger portions of space during PERT. Interestingly, mean inter-distances and convex hull of both the head and the bow were more variable across performers in NORM than in PERT (see figure 3). Individual motor strategies were thus sparser in NORM and more commonly shared in PERT. In sum, movement amplitude and dispersion of the head were smaller in PERT than in NORM, but they did not vary across conditions for the bow.

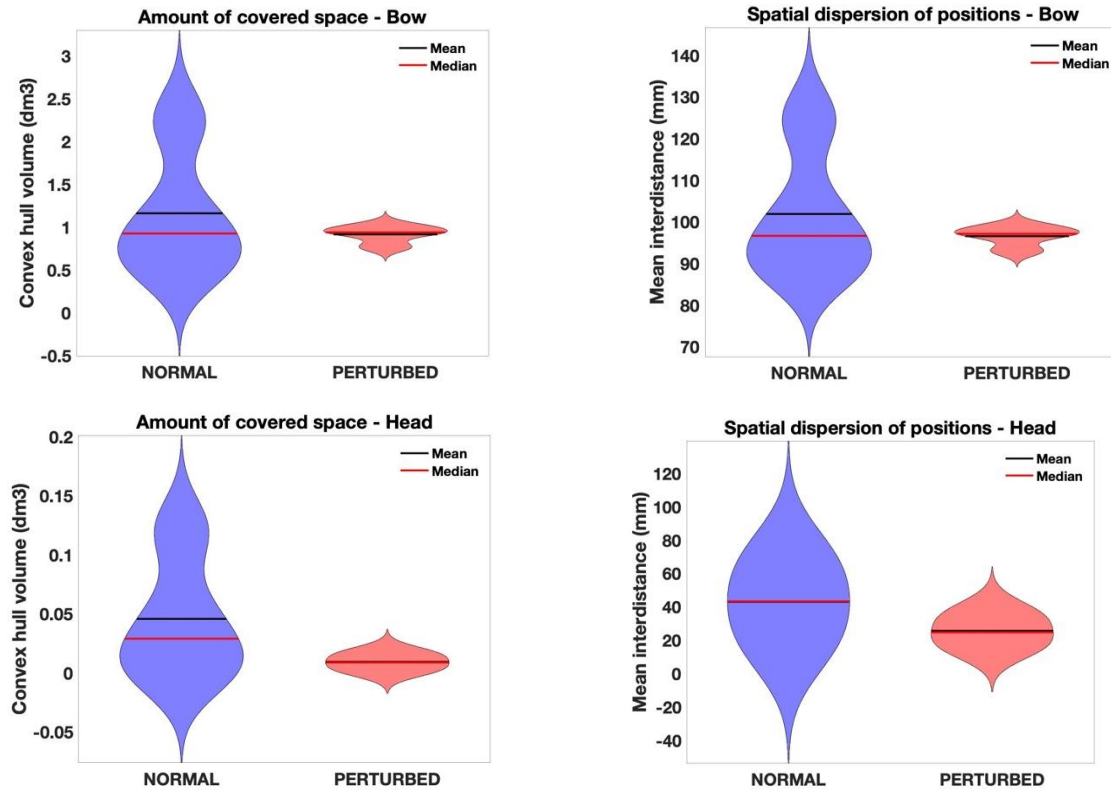


Figure 3: Violin plots of the amount of covered space (convex hull volume - left panels) and spatial dispersion (mean inter-distance - right panels) of bow (upper panels) and head (lower panels) positions in each experimental condition. Head (but not bow) motion was drastically reduced during perturbation. This partly explains the reduced variability in that condition, although it might also reflect more stable and shared movement strategies. In effect, while these variables didn't differ in magnitude across conditions for the bow motion, variability among players was also drastically reduced during perturbation.

	Mean Inter-distance		Convex Hull	
	Bow	Head	Bow	Head
NORM mean (std)	101.86 (16.11)	43.00 (24.15)	1162.69 (764.14)	45.727 (53.95)
PERT mean (std)	96.52 (2.52)	25.62 (10.15)	916.47 (106.23)	9.308 (7.27)
Chi-2	2.9127e-30	23.08	0.06	25.44
P-value	1.00	0.000002	0.8102	0.0000005

Table 1: Mean inter-distance (in mm) between positional datapoint and convex hull volume (in cm³) indicating head and bow trajectories, compared across experimental conditions. P-values in bold indicate significant differences.

Power spectral density analysis

The average spectrum of the bow velocity showed multiple peaks around 1, 2, 4 and 8 Hz (peaks observed at these frequency components will henceforth be designated as P1, P2, P3 and P4 respectively; see Figure 4a). Those frequencies match well the rates at which the rhythmical values of the score were performed: according to our audio tracks segmentation, whole, half, quarter, and eighth notes represented, on average, periods of 0.96, 1.93, 3.85 and 7.7 Hz

respectively. Beside approximations due to the width of the bins in our frequency analysis, tempo fluctuations within the piece can bring out small differences between frequency peaks in motion time series and the frequency associated to the performed rhythmical values. In particular, the last portion of the performance, which contained many of the eighth notes, was played at a faster tempo, as instructed by the score. Still, the overall spectral composition of bow motion reflected the rhythmical organization of the performance well, with multiple frequency components corresponding to the different metrical timescales of the score.

To compare frequency peaks across conditions, we extracted power at the frequency bins corresponding to each of the 4 metrical timescales described above. No difference was found across conditions for any of the peaks in the bow motion (table 2). In other words, the bow movements closely mirrored the metrical organization of the score rather than being affected by the nature of interpersonal coupling.

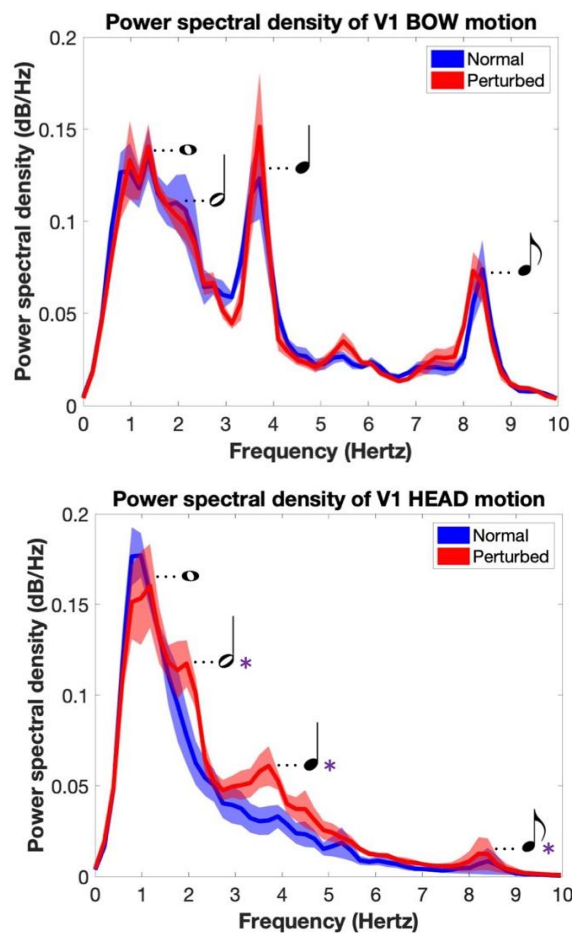


Figure 4: Power spectral density of the motion of the bow (left) and the head (right) of V1 for each experimental condition (shaded areas represent inter-subject standard error and violet asterisks denote significantly differ power across conditions). Main frequency peaks are associated with the rhythmical values they corresponded to in the performed piece (and represented here as musical notations). Bow motion was characterized by multiple periodicities that reflected the metrical organization of the score. Head motion shift from simple patterns dominated by one frequency peak near 1 Hz in NORM to more complex patterns of motion in PERT, where multiple, faster periodicities appeared and matched those found in the bow motion.

VIOLIN	P1	P2	P3	P4
NORM mean (std)	0.125 (0.031)	0.109 (0.027)	0.152 (0.065)	0.079 (0.026)
PERT mean (std)	0.133 (0.043)	0.108 (0.013)	0.147 (0.067)	0.094 (0.022)
Chi-2	1.64	0.16	0.03	2.83
P-value	0.2002	0.6889	0.8728	0.0927

Table 2. Power values (in db/Hz) extracted from the power spectral density spectra at the 4 frequency bins that corresponded to musical metrical levels (P1 - P4), compared across experimental conditions for the motion of the bow.

In contrast to bow movements, head movements clearly differ between NORM and PERT. In NORM, the frequency composition of head movements was much simpler than for bow movements (figure 4). The spectrum was dominated by a unique component situated around 1 Hz on average (P1), the frequency bin that corresponded to the average periodicity of the bar, or whole note, during the performances. During PERT, this peak persisted (although peak frequency shifted from 1.0 Hz to 1.2 Hz), but other peaks appeared around 2.0 Hz (P2), 3.7 Hz (P3), and 8.2 Hz (P4). Just like those of the bow motion, these additional peaks matched the average frequency of the metrical timescales of the performed piece well (half, quarter, eighth notes). Power observed at these additional peaks (i.e., P2, P3, P4, but not P1) was significantly higher during PERT compared to NORM (table 3). During PERT, head motion thus became more complex, displaying activity at multiple and faster metrical timescales than during NORM. Further, in PERT more than in NORM, the overall frequency composition of head motion resembled that of the bow motion.

HEAD	P1	P2	P3	P4
NORM mean (std)	0.173 (0.024)	0.096 (0.047)	0.033 (0.017)	0.007 (0.010)
PERT mean (std)	0.153 (0.051)	0.117 (0.033)	0.060 (0.014)	0.013 (0.018)
Chi-2	2.83	5.77	14.16	12.98
P-value	0.0927	0.0163	0.0002	0.0003

Table 3: Power values (in db/Hz) extracted from the power spectral density spectra at the 4 frequency bins that corresponded to musical metrical levels (P1 - P4), compared across experimental conditions for the motion of the head. Bolded p-values indicate significance ($p < .05$).

Power correlation analysis

To verify if the frequency composition of head motion during PERT reflected the mimicking of the bow rhythmical movements, we computed the correlation between power observed in head and bow of V1 at each relevant metrical timescale. Correlation coefficients were very small, and no significant difference was found between conditions for any of the timescales (Table 4). Whereas multiscale head motion patterns reflected the overall metrical structure of the score more during PERT, they did not match to the momentary rhythmical variations of the bow.

	P1	P2	P3	P4
NORM mean (std)	-0.01 (0.13)	-0.1 (0.24)	0.06 (0.16)	0.44 (0.25)
PERT mean (std)	-0.03 (0.08)	0.03 (0.28)	0.19 (0.28)	0.55 (0.33)
Chi-2	0.1	4.01	1.26	2.08
P-value	0.7488	0.0453	0.2623	0.1495

Table 4: Correlation coefficients representing the co-evolution of V1's head and bow power spectral density at the 4 frequency bins that corresponded to the piece metrical levels (P1 - P4), compared across experimental conditions.

To verify if the changes in head motion frequency composition observed during PERT were due to the visual coupling with V2, we first computed the PSD of V2's head and bow motion, and then we computed the correlation between power observed in V1's head motion and V2's head as well as bow motion at each relevant metrical timescale. In both conditions, the average spectrum of V2's bow motion mainly featured peaks close to 1 and 2 Hz (whole and half-notes, or P1 and P2), with additional small peaks around 4 Hz (quarter-note level, P3) and 6 Hz (dotted eighth notes; see figure 5). V2's head motion was dominated by a frequency component close to 1 Hz (P1), with a smaller peak around 2 Hz (P2) and a small hump around 4 Hz (P3). Beside P2, which was present in V2's motion and enhanced during PERT in V1's head motion, the frequency composition of V2's head and bow motion hardly reflected the overall changes observed in V1's head motion during PERT, where the two sections faced each other. Correlation coefficients between V1 and V2 power time series were also very small, and no difference was observed between conditions for any of the peaks (Table 5). Therefore, the evolution of the frequency composition of V1's head motion does not seem to be informed by V2's head or bow motion.

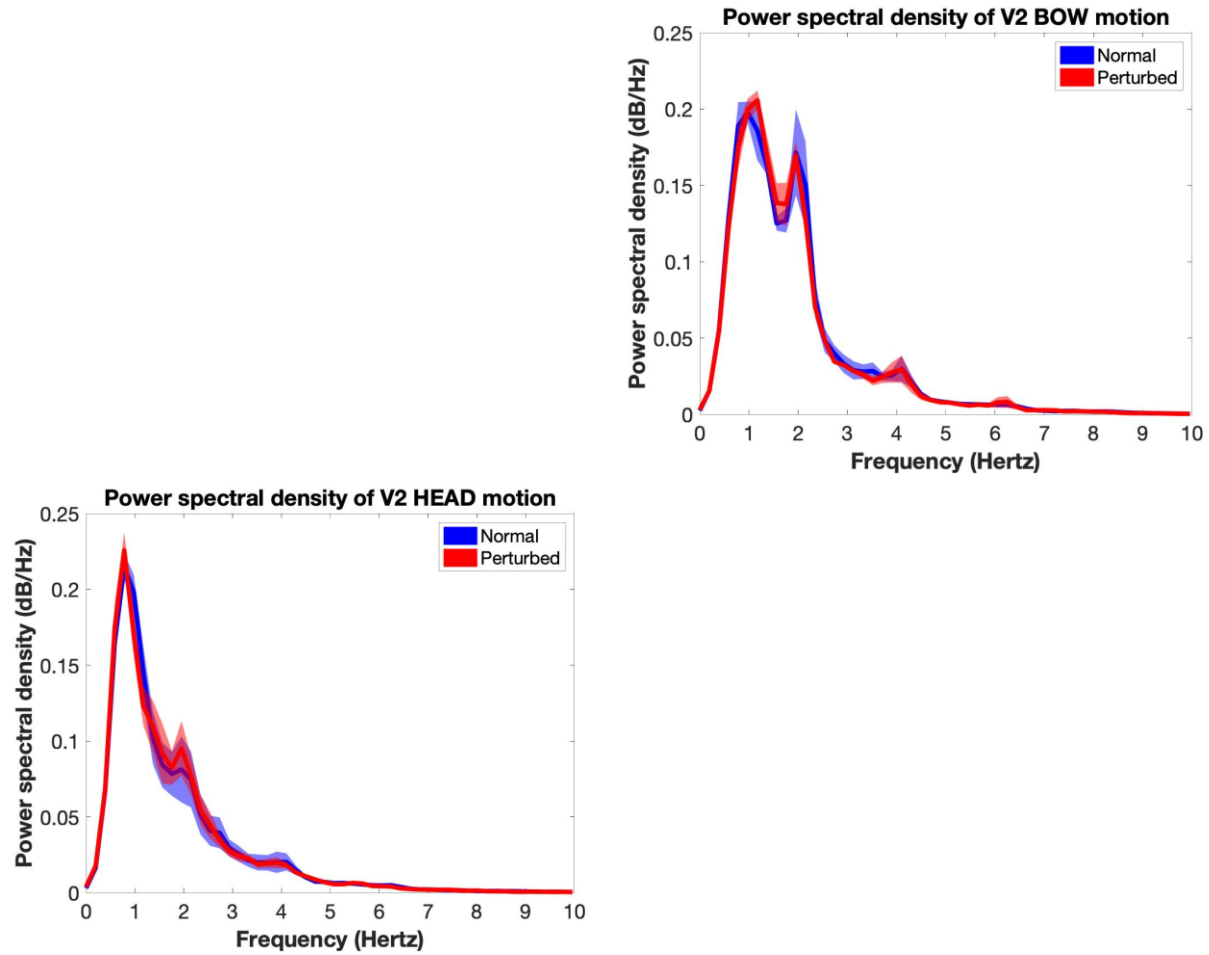


Figure 5: Power spectral density of the motion of the bow (left) and the head (right) of V2 for each experimental condition (shaded areas represent inter-subject standard error). The frequency composition of the bow motion was concentrated around 1 and 2 Hz (reflecting whole and half notes) with peaks near 4 and 8 Hz (quarter and eighth notes). The motion of the head as dominated by a component situated near 1 Hz, with an important secondary component near 2 Hz and a tiny peak near 4 Hz. No change was observed across conditions. Importantly, this frequency composition hardly explains the shift observed in V1 during perturbation, when they were facing V2.

	P1	P2	P3	P4
NORM mean (std)	0.05 (0.23)	0.08 (0.09)	0.22 (0.21)	-0.03 (0.06)
PERT mean (std)	-0.08 (0.15)	-0.09 (0.03)	-0.03 (0.11)	0.06 (0.11)
Chi-2	1.44	1.08	1.44	0.16
P-value	0.2298	0.2980	0.2298	0.6889

	P1	P2	P3	P4
NORM mean (std)	0.03 (0.23)	-0.07 (0.17)	0.18 (0.38)	0.10 (0.14)
PERT mean (std)	0.08 (0.14)	-0.10 (0.07)	0.02 (0.18)	0.30 (0.37)
Chi-2	1.26	0.03	0.1	2.56
P-value	0.2623	0.8728	0.7488	0.1093

Table 5: Correlation coefficients representing the co-evolution of V1's and V2's head (upper table) and V1's head and V2's bow (lower table) power spectral density at the 4 frequency bins that corresponded to the piece metrical levels (P1 - P4), compared across experimental conditions.

Relative phase analysis

To evaluate head and bow coupling with respect to the multiscale nature of their movements, we analyzed their phase relationships at each metrical timescale. The vector length was larger in PERT than in NORM at all timescales, and significantly so at the level of the bar (P1: Chi2 = 8.31, $p = 0.0039$) and the beat (P3: Chi2 = 5.3, $p = 0.025$; see table 6). However, the difference at P3 was only due to the trials performed with conductor 2 (see Supplementary Materials). There was also a marginal trend for the vector length to be higher in PERT than in NORM at P2 for conductor 1 only (see Supplementary Materials). The coordination between head and bow motion was therefore more stable during perturbation but at selective metrical levels only (and differentially so depending on the conductor): those metrical levels that dominated bow motion frequency composition the most. Mean phase differences were close to zero degree (i.e. in phase) for P1 and P2 and slightly negative for P3 and P4 (i.e. head motion shortly preceded the bow motion). Mean phase differences were comparable between NORM and PERT except that for P4 (Chi2 = 4.33, $p = 0.0374$; see Table 6). This indicates that, at the eighth-note level, head and bow motion were more in phase during PERT. Mean absolute phase differences were smaller in PERT than in NORM, and the difference was significant for P1, P2 and P3 (P2: Chi2 = 4.33, $p = 0.0374$; P3: Chi2 = 9.75, $p = 0.0018$; P4: Chi2 = 10.78 $p = 0.001$; see table 6, middle). Head and bow were thus moving more in phase during PERT at all metrical timescales except at the whole-note level (P1).

VL	P1	P2	P3	P4
NORM mean (std)	0.473 (0.039)	0.493 (0.057)	0.482 (0.023)	0.514 (0.076)
PERT mean (std)	0.517 (0.046)	0.518 (0.037)	0.521 (0.014)	0.537 (0.091)
Chi-2	8.31	1.85	5.3	1.85
P-value	0.0039	0.1735	0.025	0.1735

MA	P1	P2	P3	P4
NORM mean (std)	0.022 (0.258)	0.011 (1.350)	-0.331 (0.560)	-0.937 (1.088)
PERT mean (std)	-0.102 (0.343)	0.098 (0.931)	-0.223 (0.448)	-0.323 (0.866)
Chi-2	1.85	2.56	0.01	4.33
P-value	0.1735	0.1093	0.9362	0.0374

AMA	P1	P2	P3	P4
NORM mean (std)	1.190 (0.062)	1.411 (0.141)	1.405 (0.109)	1.539 (0.090)
PERT mean (std)	1.123 (0.112)	1.300 (0.160)	1.246 (0.106)	1.352 (0.179)
Chi-2	2.08	4.33	9.75	10.78
P-value	0.1495	0.0374	0.0018	0.001

Table 6: Vector length (upper table), mean relative angle (in radians, middle table) and mean absolute angle (in radians, lower table) of the relative phase between head and bow motion, computed at each metrical timescale (P1 - P4), compared across experimental conditions.

In sum, we observed three main effects in first-section violinists: 1) as expected, the overall intrapersonal coordination of head and bow motion increased, 2) qualitative shifts occurred in head (but not bow) movements: they diminished in amplitude but increased in spectral complexity, to reflect more closely the metrical structure of the score, and 3) the intrapersonal coordination of head and bow movements increased differentially at multiple timescales.

References

- D'Ausilio, A., Badino, L., Li, Y., Tokay, S., Craighero, L., Canto, R., Aloimonos, Y., & Fadiga, L. (2012). Leadership in orchestra emerges from the causal relationships of movement kinematics. *PLoS one*, 7(5), e35757.
- Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R., & Schmidt, R. C. (2007). Rocking together: Dynamics of intentional and unintentional interpersonal coordination. *Human movement science*, 26(6), 867-891.
- Riley, M. A., Richardson, M., Shockley, K., & Ramenzoni, V. C. (2011). Interpersonal synergies. *Frontiers in psychology*, 2, 38.

Schmidt, R. C., & Richardson, M. J. (2008). Dynamics of interpersonal coordination. In *Coordination: Neural, behavioral and social dynamics* (pp. 281-308). Springer, Berlin, Heidelberg.

Volpe, G., D'Ausilio, A., Badino, L., Camurri, A., & Fadiga, L. (2016). Measuring social interaction in music ensembles. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1693), 20150377.

2.3.5 Tracking the leader: gaze behaviour in group interactions

For a full description please see: Capozzi F., Beyan, C., Pierro, A., Koul, A., Murino, V., Livi, S., Bayliss, A. P., Ristic, J., Becchio, C. (2019). Tracking the Leader: Gaze Behavior in Group Interactions. *iScience*;16:242-249. doi:10.1016/j.isci.2019.05.035

It is commonly believed that leadership is reflected in visual behavior. However, little is known about how leadership shapes gaze dynamics during real-world interactions (Capozzi and Ristic, 2018; Koski et al., 2015; Risko et al., 2016). One major reason is the lack of tools to study the gaze behaviours of multiple agents in unconstrained settings.

In this study, we developed a novel tripartite method combining A) computer vision methods for remote gaze-tracking, B) a detailed taxonomy to encode the implicit semantics of multi-party gaze features, and C) advance machine learning methods to establish statistical dependencies between leadership and group visual behaviour during group discussion (Figure 1).

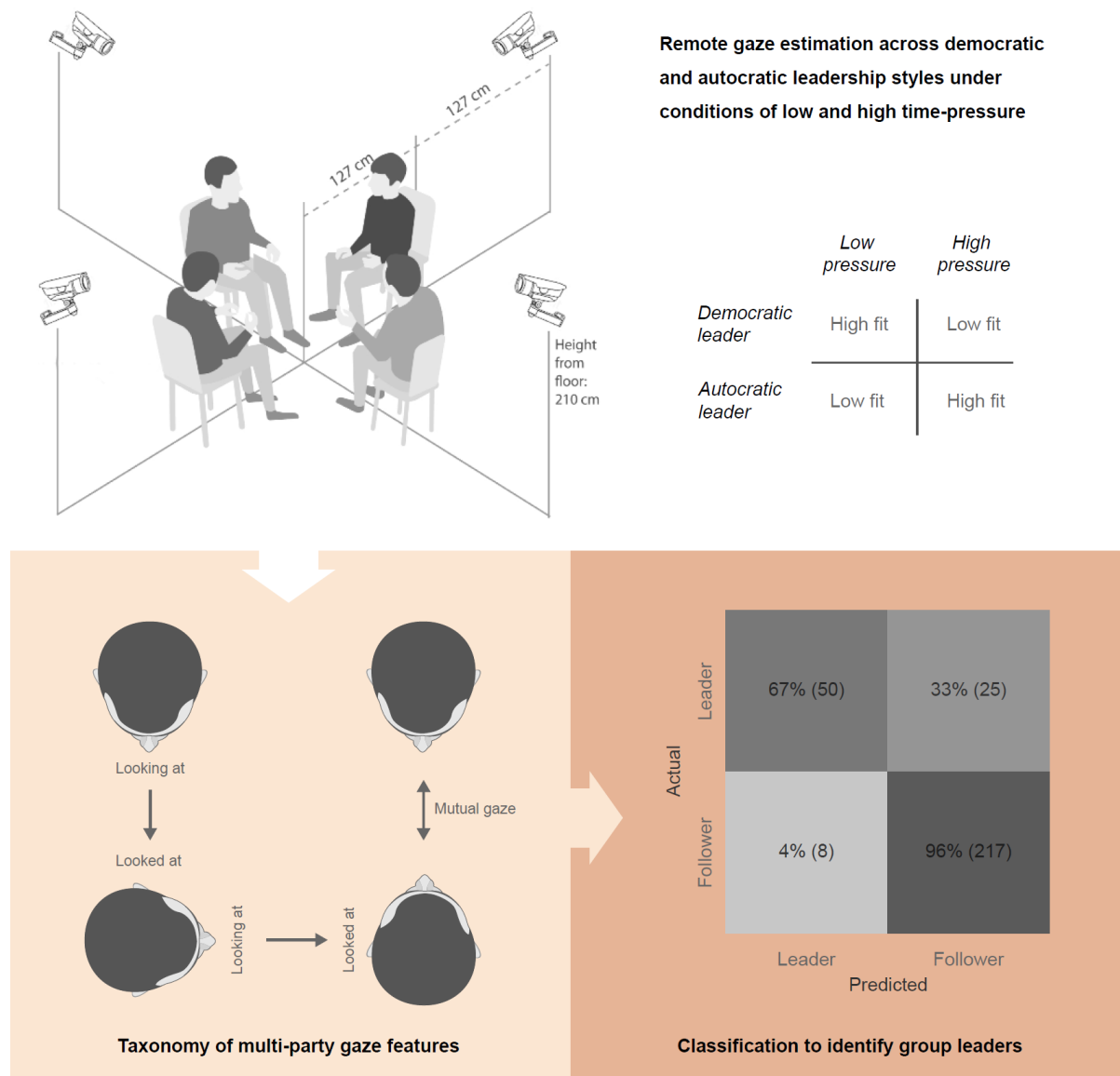


Figure 1: Depiction of gaze estimation methods applied in the study and result matrix.

The basic idea of our approach was to conceptualize multi-party gaze features as patterns and to treat the analysis as a pattern classification problem: can the semantics of group visual behaviour reveal the leader among group members? This is the first question we addressed in the study described here. The second question is whether the relationship between gaze behaviour and leadership generalizes across leadership styles and situational conditions – in other words, whether gaze behaviour can serve as a general marker of leadership.

Drawing on ideas from social psychology, we analysed gaze-based interaction dynamics in four leadership settings resulting from the orthogonal manipulation of leadership style (i.e., Democratic vs. Autocratic) and situational condition (i.e., Low time-pressure vs. High time-pressure). Democratic leadership is expected to be more effective under situational conditions of low time-

pressure, whereas autocratic leaderships is expected to be more effective under situational conditions of high time-pressure. The orthogonal manipulation of leadership styles and situational conditions resulted in two high-fit conditions (Democratic - Low time-pressure, Autocratic - High time-pressure) and two low-fit conditions (Democratic - High time-pressure, Autocratic - Low time-pressure) (Figure 6). Each group, composed of one designated leader and three followers, was assigned a survival task to solve within a limited time.

First, using a method for automatically estimating the Visual Focus of Attention (VFOA), we determined 'who looked at whom'. Then, we established a detailed taxonomy of multi-party gaze behaviours and, combining the VFOA of individual group-members, reconstructed the gaze-based interaction dynamics. Next, we probed the actual association between leadership and gaze patterns by asking whether a pattern classification algorithm (SVM) could discriminate between leaders and followers among the group-members (Koul et al., 2018).

Our approach was successful and extremely revealing. We found that social gaze behaviour distinctively identified group leaders. With a cross-validated accuracy of 89%, classification performance was well above the .50 chance level (95% CI = .85, .92; Kappa = .68; Sensitivity = .86; Specificity = .90; F1 = .75; $p < .001$) (Figure 6). To investigate which features were more effective for the classification task, we next computed F-scores (see "Leader classification analysis" in Transparent Methods). F-score provides a measure of how well a single feature at a time can discriminate between different classes. The higher the F-score, the greater the ability of a feature to discriminate between leaders and followers. Table 1 provides an overall view of the discriminative power of each visual feature. Overall, F-scores suggest that leaders looked less at others and, conversely, were looked at more as compared with followers. Also, leaders were involved in and caused more episodes of mutual engagement, relative to followers. The time taken by another group member to respond to the initiation of mutual engagement was also less for leader-initiated episodes compared to follower-initiated episodes.

Table 1

F-Scores and Group Means for Individual Features for Discrimination between Leaders and Followers (Full Dataset)

Feature	F-Score	Leaders Mean ($\pm$ SD)	Followers Mean ($\pm$ SD)
Looking at	1.800	0.36 $\pm$ 0.09	0.57 $\pm$ 0.13
Looked at_Ratio	1.700	2.43 $\pm$ 1.07	0.85 $\pm$ 0.53
Looked at	1.300	0.72 $\pm$ 0.18	0.43 $\pm$ 0.17
Looked at_multiple	1.300	0.28 $\pm$ 0.13	0.10 $\pm$ 0.08
Mutual gaze	0.780	0.41 $\pm$ 0.14	0.24 $\pm$ 0.12
Mutual gaze_mutiple	0.450	0.26 $\pm$ 0.14	0.15 $\pm$ 0.10
Mutual gaze response time	0.350	0.13 $\pm$ 0.06	0.19 $\pm$ 0.08
Mutual gaze initiation	0.085	0.27 $\pm$ 0.08	0.24 $\pm$ 0.07

test-democratic led to a similar cross-classification accuracy of 90% (95% CI = .84, .95; Kappa = .72; Sensitivity = .89; Specificity = .91; F1 = .78; $p < .001$).

With a similar logic, we applied MVCC to test generalization across situational conditions.

We trained a linear SVM on gaze patterns recorded under high fit situational conditions (i.e., democratic leaders working in a low time-pressure condition and autocratic leaders working in a high time-pressure condition), and then tested it on group interactions under low fit situational conditions and vice-versa. Cross-classification performance was once again well above the .50 chance level, reaching 94% and 85% for train-high fit and test-low fit (95% CI = .89, .97; Kappa = .83; Sensitivity = .92; Specificity = .94; F1 = .87; $p < .001$) and train-low fit and test-high fit (95% CI = .78, .91; Kappa = .54; Sensitivity = .82; Specificity = .86; F1 = .63; $p < .001$), respectively. Collectively, these data show that multi-party visual behaviour supports identification of group leaders across leadership styles (i.e., democratic, autocratic) and situational fit conditions (i.e., high fit, low fit).

To our knowledge, this is the first study that attempts to provide a full characterization of the relationship between leadership and social gaze behaviour during natural group interactions. The novel method utilized in the current study demonstrates that gaze-based group behaviours distinctively identified leaders during natural group interactions. Leaders were looked at more, looked less at others, and elicited more mutual gaze. This pattern was observed over time regardless of leadership style and situational condition, suggesting that gaze can serve as a *general marker of leadership*. Together with previous findings on body movements and paralinguistic behaviours, these results demonstrate the significance of non-verbal cues for leadership identification. We expect that future empirical and modelling studies will investigate whether and how different (and possibly correlated) non-verbal features contribute to leader classification.

References

- Capozzi, F., Ristic, J., 2018. How attention gates social interactions. *Ann. N. Y. Acad. Sci.* 1426, 179–198.
- Koul, A., Becchio, C., Cavallo, A., 2018. PredPsych: A toolbox for predictive machine learning-based approach in experimental psychology research. *Behav. Res. Methods* 50, 1657–1672.
- Koski, J.E., Xie, H., Olson, I.R., 2015. Understanding social hierarchies: The neural and psychological foundations of status perception. *Soc. Neurosci.* 10, 527–550.
- Risko, E.F., Richardson, D.C., Kingstone, A., 2016. Breaking the fourth wall of cognitive science: Real-world social attention and the dual function of gaze. *Curr. Dir. Psychol. Sci.* 25, 70–74.

2.3.6* Capturing human movement and shape information from small groups to extract expressive and social features - using marker-less techniques

*Study is not originally listed in the D1.2

Human beings exhibit phenomenal capabilities in synchronizing joint actions and coordinating at the inter-personal level in a non-verbal manner. This is observed specifically in musical ensembles where co-performers are seen to coordinate their movements effortlessly (Bishop, 2018). Perhaps the most natural response to music is to move and synchronize to the rhythmic elements in and inter-twined in music. When listening to music, we tend to raise our hands, tap our feet, dance, and shake our heads. It has been observed in a musical ensemble, that when a musical piece is being played, there are parts or phrases, basis which the members tend to coordinate their movements with the rhythmic behavior of other group members.

We, as individuals, usually have a general feeling of entrainment, but almost unknowingly, yet spontaneously, we move to music being played around us (Bishop & Goebel 2018a). We may react genuinely to music, and sometimes showcase unique movements in response to what is being heard (Luck et al. 2010; Vuoskoski et al. 2011). The mix of music and the corresponding movement seems to trigger a social bonding effect. These movements or bodily gestures tend to convey certain subtle messages, and this conveyance of messages is crucial for the co-creation of a musical piece, and musicians are seen to perpetually move during a performance to augment the creation of sound, express themselves, communicate with their fellow group members, and transition into states of synchronization (Wanderley, 2001; Bishop & Goebel 2018b). This coordination often takes place in the context of different musical texture that vary in terms of whether there is a clear hierarchy with a leader playing the melody while others play the accompaniment.

Music permits the study of these subtle facets of the human body, using current state-of-the-art methods, while producing minimal noise during experiments. This noise is the random variability one may find in a signal. One way to study such an area is extracting the signals that emerge out of human movements. But when one must perform experiments in-the-wild, noise is a major concern – which adds to the numerical complexities. Another benefit of experimenting with music is that it allows research to be carried out with more care, attention to detail, and control. Music ensemble performances are in turn special examples of joint actions with key advantages. In our experiments, we analyze each of the videos in phrases. These phrases are units of information at relatively long musical timescales. Interpersonal coordination measures using a windowing approach captures shorter timescales. During our analysis we examine different positions of the phrase (start, middle, end). With this, we explore the opportunity to answer questions related to multiple timescales – especially how coordination at shorter timescales changes over the course of longer timescales.

We propose a computational model to compute the synchronization of dyadic pairs in a musical ensemble using marker-less computer vision techniques. Our methods involve the use of human pose estimation algorithms. The human body in pose estimation algorithms is looked at as a system with many elements. Each of these elements, called a key-point, is tracked in a consistent manner through the visual sequences. Eventually, on discovering the position of a group of anatomical joints such as elbows, nose, shoulders, knees, etc., these key-points help identify a blue- print where a skeleton-like structure of the body can be super-imposed. The algorithm then proceeds with completing a pair-wise connection between these key-points to provide a human skeletal structure, super-imposed on the input, known as a Pose. On implementing these algorithms on our dataset, we receive an output as a json file which contains the coordinates of essential body joints.

Interestingly, in a musical ensemble, due to being seated, musicians communicate with each other not using speech, but head and other upper-body movements. They can interact with each other with visual or audio cues. This interaction usually conveys a message to initiate a musical piece at a certain time and could also convey how musical notes need to be played. Thus, we focus on investigating techniques for the automatic analysis of synchronization by tracking the movement of the human head (Yokozuka et al. 2018). The head is tracked by making use of the coordinates of the human nose as provided by the pose estimation algorithm. The kinematic information extracted from the head movements of the performers are then subject to certain signal processing techniques which also find use in neuroscientific research activities to assess connectivity and synchronization of different regions in the brains. We use the data to compute dyadic synchronization between the performers using a metric called the Phase-Locking Value.

The dataset used for these experiments consist of videos from concert performances by the Omega Ensemble, a professional chamber music group from Australia. Each of the videos have been annotated as determined by a musicological analysis based on the published score. For our research the concert video on which experiments were performed was “Brahms Clarinet Quintet” in B minor (Op.115) written in 1891. This work has three movements, which are each designated as a “piece” in the analysis below. Each of these videos being experimented upon have 5 participants.

The goal of our study is to investigate factors that influence the quality of interpersonal coordination between group members: 1. Strength; and 2. Directionality of interpersonal coupling in musical ensemble performance.



Figure 1: Image from a musical piece composed by Johannes Brahms

The specific aim of this analysis is to test how strength and directionality of coupling are influenced by two factors: 1. Position within the musical phrase (Start, Middle, and End); and 2. Musical textures (Homophonic and Polyphonic).

Previous research (Schögler, 1999; Keller et al. 2014) suggests that coupling may be stronger at the beginning and end of phrases than in the middle. The expectation is that when leadership is not assigned there is better interpersonal coordination. This would suggest that coupling strength will be stronger when leadership is distributed – stronger in polyphonic than homophonic texture conditions (Novembre et al. 2015; Noy et al. 2011; Varlet et al. 2020). Also, the presence of a leader may be more influential at the beginning and ending of phrases than in the middle, in which case we would expect a statistical interaction of the two factors. Our study is strongly based on agreements that performers in a musical ensemble produce more coordinated gestures, and by finding the dyadic synchronizations using performer head – movements, we get closer to answering some of these open-ended questions (King et al. 2011). Using non-intrusive techniques, we also present a new computational method to test and present results on the above hypothesis.

The inter-personal coordination is observed by obtaining the dyadic synchronization on selecting a dyad combination of two co-performers in a musical group. For a set of 5 participants (Figure 1), we can have a total of 10 combinations, and a total of 26 video samples have been studied. This includes deciding within the pairs who shall be the melody and accompaniment for between subject effects. This was established on the basis of the musicological analysis, which classified whether each part was playing the melody or accompaniment in the musical score. We then decided that texture will be a between subjects' factor in the Repeated Measures ANOVA

because we did not have control over the textures since they were specified in the musical score. Also, no two people are playing the same part – thus also permitting the inclusion of factors such as instruments and understanding their individual effects on a musical phrase.

Our results investigate and demonstrate the variation in body movements during different kinds of musical textures, and delicate transmission of messaging at multiple temporal scales. The Phase Locking Values (PLV) for all pairs were entered into Repeated Measures ANOVA to test for effects of position in the phrase, texture, and pair.

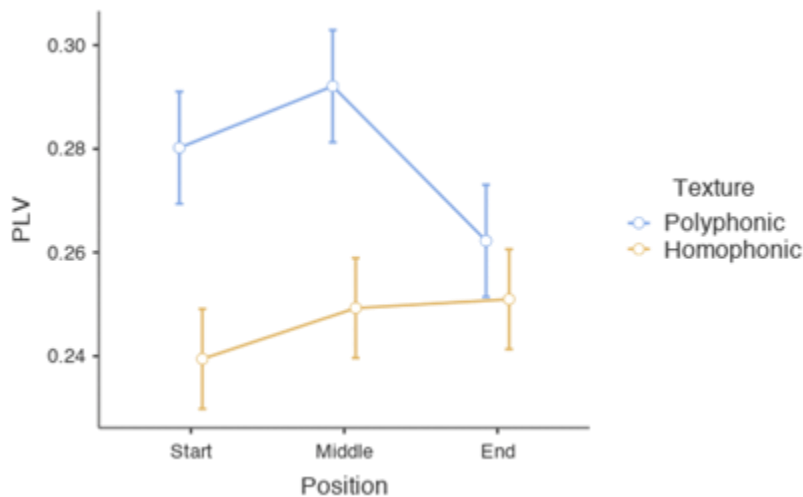


Figure 2: Measure output for Position x Texture

From Figure 2, we observe how the PLV begins at a lower value in both textures, Polyphonic and Homophonic. These PLVs while they start out higher, it tends to rise until the middle of the phrase, and then begins to drop in value towards the end of a musical phrase. It is typically seen from the data plotted that those polyphonic textures have stronger coupling than homophonic textures.

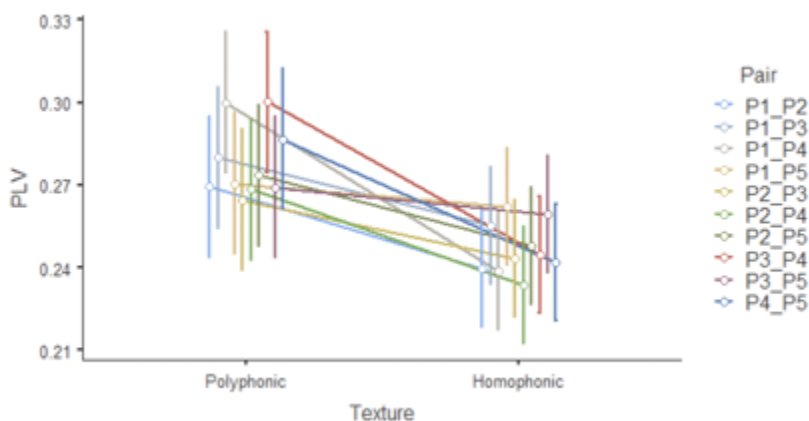
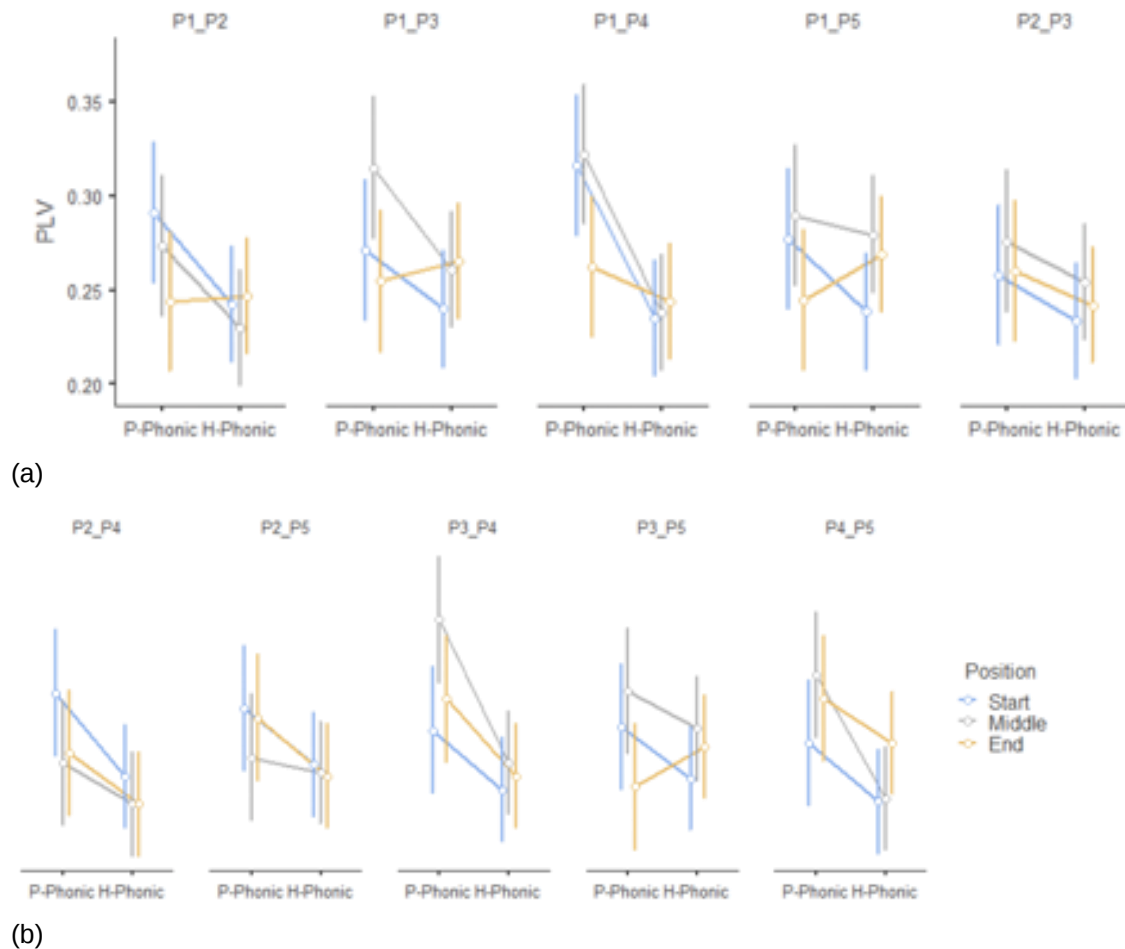


Figure 3: Measure output for Texture x Pair

There were 5 players that synchronized their notes in both the homophonic and polyphonic textures. As seen in Figure 3, pertaining to a pair – wise analysis, we notice that all pairs show a higher level of synchronization in polyphonic textures. We see a sharp or marginal drop in all pairs of performers, suggesting that there may exist sufficient conditions to initiate a leader and follower relationship. This could be because the four performers in a homophonic texture are coupled to the melody player, whereas in a polyphonic texture the coupling is evenly distributed across all performers.

**Figure 4:** Measure output for Texture x Position x Pair

In Figure 4, P-Phonic stands for polyphonic while H-Phonic stands for homophonic, and the results have been arranged in a pairwise order. They depict the results of marginal means of the position as well. We observe that nearly all pairs are seen to have higher PLV values in the middle section of the phrases. In some examples such as P2_5, P2_4, and P1_2, this may differ because while there may be head movements involved, our video is being captured from the front. Thus, the ones seated on the extreme left and extreme right will exhibit more numerical movement due to being seated in a different orientation which can capture better trajectories. Additionally, the

drop in Phase Locking Values at the end of the phrase for polyphonic textures could indicate that one of the instruments takes over as leader at that point, thus making it like a homophonic texture and reducing the symmetry of coupling across the ensemble.

We found statistically significant main effects of Position, $F(2, 400) = 4.657$, $p=0.01$, and Texture, $F(1, 200) = 34.689$, $p<0.001$. A significant effect of Piece was also observed, $F(2, 200) = 54.500$, $p<0.001$, but this is currently beyond our theoretical interest. Additionally, the two-way interaction between Position and Texture is statistically significant, $F(2, 400) = 6.658$, $p=0.001$. This indicates that there is a matching effect between the Position and Texture in musical phrases that have been analyzed – which is in line with the specific aim of our hypothesis and confirms the reliability of the overall pattern of results described in the sections above. Thus, based on our results, we can say that the effect of the texture changes with respect to the position of the phrase, further helping us study the relationship between musical textures and overall coupling strengths of performers over the course of a musical phrase. In continuation to the previously reported results in D2.5, we also performed an additional analysis to examine the relationship between the coordination of body motion, which provides visual cues, and the synchrony of ensemble sounds.

Because we do not have multitrack audio recordings for each instrument on a separate track, we computed indirect measures of global ensemble synchrony from stereo auditory recordings of the full ensemble sound. Based on previous research (Clayton et al., 2020; Eerola et al., 2018; Jakubowski et al., 2020), we included estimates of 'pulse clarity' and 'event density', which were calculated using the 'mirpulseclarity' and 'mireventdensity' functions from the MIRtoolbox in MATLAB (Lartillot, Toivianen, & Eerola, 2008). Pulse clarity is a feature that reflects the strength of rhythmic beats, while event density indicates the average frequency of events (i.e., the number of events detected per second). Descriptive statistics for these measures are shown in Table 1.

To assess the potential effects related to these audio features, we ran a linear mixed effects model analysis using the *lmer* package (Bates, Maechler, Bolker, & Walker, 2014) in R (R Core Team, 2020) with PLV as the dependent variable, pulse clarity, event density, texture, and phrase position as predictor fixed effects, and piece as a random effect (with intercepts allowed to vary). Pulse clarity values were arcsin-transformed and event density values were log-transformed prior to analysis. The results revealed a link between PLV and event density. Specifically, a likelihood-ratio test indicated that a model including event density provided a better fit for the data than a model without it ($\chi^2(1) = 7.44$, $p = .006$), whereas pulse clarity did not contribute significantly to the model ($\chi^2(1) = 0.03$, $p = .884$). Examination of the output for the full model indicated that PLV values increased with increasing event density ($\beta = 0.031$, $SE = 0.011$, $t = 2.767$, $p = .006$).

These results are consistent with a growing body of evidence that visual and audio cues are both relevant in assessing interpersonal synchrony in musical ensembles (Clayton et al., 2020; Eerola et al., 2018; Keller & Appel, 2010; Jakubowski et al., 2020; Ragert et al., 2013). Future work with multitrack audio would allow the relationship between auditory and visual information to be investigated in greater detail, including the assessment of correspondence between leader-follower relations across modalities.

		Texture					
		Homophonic			Polyphonic		
		Phrase Position					
Piece	Measure	Start	Middle	End	Start	Middle	End
Brahms	Pulse Clarity						
	Mean	0.111	0.122	0.167	0.148	0.126	0.135
	SD	0.046	0.057	0.058	0.075	0.047	0.052
	Event Density						
	Mean	1.416	2.098	1.767	1.733	2.105	2.205
	SD	0.844	1.484	0.931	0.846	0.819	0.845
Borodin	Pulse Clarity						
	Mean	0.153	0.157	0.152	0.16	0.156	0.144
	SD	0.063	0.044	0.083	0.081	0.054	0.058
	Event Density						
	Mean	2.102	2.202	1.631	2.197	2.025	2.033
	SD	0.946	0.934	0.799	0.798	0.792	1.25

Table 1: Mean and standard deviation (SD) of estimates of pulse clarity and event density as a function of texture (homophonic and polyphonic) and phrase position (start, middle, and end) for performances of pieces by Brahms and Borodin.

In the future, we will be exploring various techniques to examine how interpersonal coordination unfolds at multiple timescales – and this could involve applying these techniques in different experimental setups. They could include:

- Cross-spectral coherence where computation of synchrony is across a range of timescales reflecting coordination of different body parts (Richardson et al., 2005, Varlet et al., 2015)
- Cross-wavelet coherence where we will be studying datasets with multiple frequencies involved (Walton et al., 2015)
- Multiscale Entropy if we are dealing with unknown or skewed data distribution (Glowinski et al., 2010)
- Multi-Event Class Synchronization if we have discreet information to help us measure synchronization between two relevant events that belong to different event classes and detected in multiple time series (Alborno et al., 2019); and
- Granger Causality to quantify mutual influence/leadership by studying the directionality of coupling, helping us look at effects of musical structure on group coordination and communication simultaneously, at short timescales related to musical beats and longer timescales related to expressive body sway (Hilt et al., 2019).

References

Alborno, P., Volpe, G., Mancini, M., Niewiadomski, R., Piana, S., & Camurri, A. (2019). The Multi-Event-Class Synchronization (MECS) Algorithm. arXiv preprint arXiv:1903.09530.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823

Bishop, L. (2018) "Collaborative musical creativity: How ensembles coordinate spontaneity," *Frontiers in psychology*, vol. 9, p. 1285.

Bishop, L. & Goebel, W. (2018a) "Beating time: How ensemble musicians' cueing gestures communicate beat position and tempo," *Psychology of music* (46)1, 84– 106.

Bishop, L. & Goebel, W. (2018b) "Communication for coordination: gesture kinematics and conventionality affect synchronization success in piano duos," *Psychological Research*, vol. 82, no. 6, 1177–1194.

Clayton, M., Jakubowski, K., Eerola, T. , Keller, P. E., Camurri, A., Volpe, G. and Alborno, P. (2020). "Interpersonal entrainment in music performance: Theory, method, and model," *Music Perception: An Interdisciplinary Journal*, vol. 38, no. 2, pp. 136–194, 2020.

Eerola, T., Jakubowski, K., Moran, N., Keller, P. E. and Clayton, M. (2018). "Shared periodic performer movements coordinate interactions in duo improvisations," *Royal Society open science*, vol. 5, no. 2, p.171520.

Glowinski, D., Coletta, P., Volpe, G., Camurri, A., Chiorri, C., & Schenone, A. (2010, October). Multi-scale entropy analysis of dominance in social creative activities. In *Proceedings of the 18th ACM international conference on Multimedia* (pp. 1035-1038).

Lartillot, O., Toiviainen, P., & Eerola, T. (2008). A Matlab toolbox for music information retrieval. In C. Preisach, H. Burkhardt, L. Schmidt-Thieme, & R. Decker (Eds.), *Data analysis, machine learning and applications. Studies in classification, data analysis, and knowledge organization* (pp. 261–268). Berlin, Heidelberg: Springer. http://dx.doi.org/10.1007/978-3-540-78246-9_31.

Luck, G. Saarikallio, S., Burger, B., Thompson, M., and Toiviainen, P. (2010). "Effects of the big five and musical genre on music-induced movement," *Journal of Research in Personality*, vol. 44, no. 6, 714–720.

Jakubowski, K., Eerola, T., Blackwood Ximenes, A., Ma, W. K., Clayton, M., & Keller, P. E. (2020). Multimodal perception of interpersonal synchrony: Evidence from global and continuous ratings of improvised musical duo performances. *Psychomusicology: Music, Mind, and Brain*, 30(4), 159–177. <https://doi.org/10.1037/pmu0000264>.

Keller, P. E., & Appel, M. (2010). Individual differences, auditory imagery, and the coordination of body movements and sounds in musical ensembles. *Music Perception*, 28(1), 27-46.

Keller, P., Novembre, G. & Hove, M. (2014) "Rhythm in joint action: psychological and neurophysiological mechanisms for real-time interpersonal coordination," *Philosophical Transactions of the Royal Society B: Biological Sciences* (369)1658, 20130394.

King, E., & Ginsborg, J. (2011) "Gestures and glances: Interactions in ensemble rehearsal," *New perspectives on music and gesture*, 177–2011.

Novembre, G., Varlet, M., Muawiyath, S., Stevens, C. J. & Keller, P. (2015) "The e-music box: an empirical method for exploring the universal capacity for musical production and for social interaction through music," *Royal Society open science* (2)11,150286.

Noy, L., Dekel, E., & Alon, U. (2011) "The mirror game as a paradigm for studying the dynamics of two people improvising motion together," *Proceedings of the National Academy of Sciences* (108) 52, 20947–20952.

Ragert, M., Schroeder, T., & Keller, P. E. (2013). Knowing too little or too much: the effects of familiarity with a co-performer's part on interpersonal coordination in musical ensembles. *Frontiers in Psychology*, 4, 368.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>

Richardson, M. J., Marsh, K. L., & Schmidt, R. C. (2005). Effects of visual and verbal interaction on unintentional interpersonal coordination. *Journal of Experimental Psychology: Human Perception and Performance*, 31(1), 62.

Schögler, B. (1999) "Studying temporal co-ordination in jazz duets," *Musicae scientiae* (3)1, 75–91.

Varlet, M., Bucci, C., Richardson, M. J., & Schmidt, R. C. (2015). Informational constraints on spontaneous visuomotor entrainment. *Human movement science*, 41, 265-281.

Varlet, M., Nozaradan, S., Nijhuis, P., & Keller, P. (2020) "Neural tracking and integration of 'self' and 'other' in improvised interpersonal coordination," *Neuro Image* (206), 116303.

Vuoskoski, J. K., Thompson, W. F., McIlwain, D. and Eerola, T (2011) "Who enjoys listening to sad music and why?," *Music Perception* (29)3, 311–317.

Walton, A. E., Richardson, M. J., Langland-Hassan, P., & Chemero, A. (2015). Improvisation and the self-organization of multiple musical bodies. *Frontiers in psychology*, 6, 313.

Wanderley, M. (2001) "Quantitative analysis of non-obvious performer gestures," in *International Gesture Workshop*, 241–253.

Yokozuka, T., Ono, E., Inoue, Y., Ogawa, K.-I. & Miyake, Y. (2018) "The relationship between head motion synchronization and empathy in unidirectional face-to-face communication," *Frontiers in psychology* 9, 1622.