This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2019.2951368, IEEE Transactions on Affective Computing

The change matters! Measuring the effect of changing the leader in joint music performances

Giovanna Varni*, Maurizio Mancini*, Luciano Fadiga, Antonio Camurri, and Gualtiero Volpe

Abstract—In a joint action, a group of individuals coordinate their movements to reach a shared goal. When a change – i.e., an event that affects group functioning – occurs, the group adopts strategies to face it. This paper investigates how a change involving a strategic core role in a group affects interpersonal coordination and ultimately group effectiveness in performing a joint action. Following the entrainment theory, interpersonal coordination is addressed in terms of the rhythmic cycles of the individuals and of the group and their adjustment. Music is used as an ideal ecological scenario for investigation. More specifically, this work focuses on orchestra playing. By adopting a computational approach, research is devoted to measure how a change of conductor (i.e., the leader) influences entrainment between players and its variation over time as well as the relationship between entrainment and external ratings of the orchestra performance. Results show that whereas the change of conductor had a limited significant effect on entrainment, a significant effect was found when entrainment is used as a predictor of the external ratings. Both the obtained results and the techniques developed for measuring entrainment may open novel research directions in the area of automated analysis of group behavior, and particularly of emotion in groups.

Index Terms—Joint action, group, change, entrainment, musical performance.

1 INTRODUCTION

During social interaction, multiple individuals coordinate their movements in order to communicate meaning, emotion, or to produce a change in the environment [1]. This coordination is referred as *joint action*: people coordinate along several dimensions (e.g., time and space) managing, in a spontaneous or planned way, verbal and nonverbal signals, exchanging information, and mutually adjusting their behaviors in order to achieve a common goal [2], [3]. The goal can consist in fulfilling the members' needs, communicating information to other people, and so on. For example, coordination is a key mechanism to elicit emotion in groups [4]. Several definitions of coordination exist. In group research, the one collecting most consensus addresses coordination as *"the process of orchestrating the sequence and timing of interdependent actions"* [5]. This implies that coordination is the core process through which some characteristics of the members of the group are molded to achieve the final goals.

In the last four decades, psychologists started to investigate group dynamics focusing on how information flow passes among the group's members, modeling group effectiveness, and group changes. In particular, *change* was argued as an integral component of the group's life [6]. Changes can occur in terms of membership, role, task, and external environment and they affect group functioning negatively and positively. Concretely, groups usually develop routines, that is, they elaborate preferred patterns of interaction resulting in a greater members' cohesion and in a sort of balancing. When a change occurs, this balancing is threatened and coordination strategies should be implemented to achieve the goals.

To date, two theoretical complementary frameworks were proposed to address the effect of member changes in coordination and, consequently, in group effectiveness. The first framework deals with *small groups as complex systems* [7] and focuses on the patchwork of multiple bidirectional relationships established among the members. The second framework is known as *the entrainment theory* [8] and focuses more on the effect of time: individuals and groups have their own rhythmic cycles influencing one each other. These cycles are captured by external pacers, which pull them to establish a dominant temporal ordering that becomes the coordination mechanism for all the members. Computational studies providing quantitative measures of the effects of a change are however still scarce. The advent of more and more sophisticated and ubiquitous technologies such as, for example, artificial partners conceived to collaborate with humans, suggests that a deeper investigation on techniques to measure the effects of a change in a group (e.g., the entrance of an artificial member) are needed.

In this paper, we measure how the change of a person playing a *strategic core role* [9] in a group performing a joint action affects group effectiveness. The analysis is performed by adopting a computational approach. The work focuses on orchestra playing.

Music is a joint action wherein the players continuously adapt their nonverbal behaviors to control the dynamics of the interaction. Moreover, it is a social activity recurring in all cultures and showing coordination mechanisms also detectable in other social rhythmic activities (e.g., talking and dancing). For all these reasons, it was used in this paper as an ideal ecological scenario for scientific investigation.

G. Varni is with LTCI, Télécom Paris, Institut polytechnique de Paris, France
M.Mancini is with the School of Computer Science and Information Technology, University College Cork, Ireland; he carried out the work in the paper while he was at DIBRIS, University of Genoa, Italy
G. Volpe, and A. Camurri are with DIBRIS, University of Genoa, Italy. e-mail: gualtiero.volpe@unige.it
L. Fadiga is with Italian Institute of Technology, Genoa, Italy.
**The first two authors equally contributed to the presented work and to the paper preparation.*

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2019.2951368, IEEE Transactions on Affective Computing

Referring to the second framework described above, our research is devoted to measure how a change of conductor (i.e., the leader) influences entrainment between players and its variation over time as well as the relationship between entrainment (and its variation) and external ratings reflecting orchestra effectiveness.

In musical performance, entrainment goes beyond the technical actions performed to follow a musical score: it also includes changes and adaptations to characteristics such as phrasing, tempo, and loudness, to create an aesthetically pleasant music performance, and ultimately affects the extent to which performers engage effectively in joint action [1], [10]. Moreover, several studies show that entrainment is linked with affect (e.g., [1], [10], [11], [12], see Section 2 for more details), and, in particular, it has a central role in elicitation of emotion in groups [4]. Thus, although this paper does not directly address how group entrainment affects emotion in the group itself or in an external observer, its findings may support the design of computational models of affect in groups, that is a major goal of Affective Computing [13].

According to the findings of previous studies (e.g., [12], [14], [15], [16]), we computed entrainment by measuring motor synchrony and we analyzed its variation over time. We applied an extended version of the Kuramoto model of synchrony of coupled oscillators, and a computational method for detecting entrainment patterns by observing deviations in synchrony. To assess a possible effect of the conductor change: (i) we tested whether this is detectable in the obtained measures of entrainment and its variation over time; (ii) we studied the relationship between such measures and external ratings of the orchestra performance, and we assessed whether an effect of the conductor is detectable on the model that better represents such a relationship. In order to catch the deviations from a strict synchrony between the orchestra sections, we took inspiration from [4], [14], [17] and we conceived U-shaped or Monotonic Synchrony Pattern ($UMSP$).

2 RELATED WORK

2.1 Entrainment and musical performance

As a recent survey [18] witnesses, analysis of interpersonal coordination in music ensembles is receiving a growing interest in several research communities, including musicology, psychology, neuroscience, and computer science. Musical performance involves the temporal coordination of actions between individuals in ways that are often more complex or more precise than in other social situations. According to Clayton and colleagues [19], entrainment in musical performance supports synchronization of musical sounds and higher-level musical coordination. The authors argued that at short time-scales a mechanism of sensorimotor synchronization contributes to phase alignment between sounds, whereas coordination at larger time-scales allows musicians to manage alignment of metrical and phrase structures and transitions between pieces or sections. Entrainment thus affects tempo, dynamics, metrical, and cadential structure of music. According to Phillips-Silver and Keller [1], entrainment supports the two primary generic modes of musical joint action, i.e., *chorusing* and *turn-taking*. Keller argued that "*musicians rely upon core cognitive-motor ensemble skills that are based on entrainment, and allow each individual to coordinate with the actions of co-performers in real-time to engage effectively in joint action*" [1] (see also [10]). Moreover, entrainment also influences audience's perception of a music performance. Doffman [20] suggested that musicians can use entrainment as a performative device to increase expressivity and participation. Leman [21] argued that audience and performers are linked by entrainment processes so that actions performers make influence listeners' response to music, which in turns influences performers' actions.

2.2 Entrainment and affect

The link between entrainment and affect in music has been subject of several studies. Phillips-Silver and Keller [1] argued that entrainment has two components, a temporal one and an affective one. According to them, "affective entrainment involves the formation of interpersonal bonds and is related to the pleasure in moving the body to music and being in time with others". In a study on groove, understood as "a sensorimotor phenomenon with affective consequences", Janata and colleagues [12] showed that "the quality of the sensorimotor coupling is reflected in the subjective experience" of music, thus establishing a link between the temporal and the affective component of entrainment. Leman [11] considers entrainment as "an element of expressive interaction". Juslin and colleagues [22], [23] identified rhythmic entrainment as one of the eight psychological mechanisms that explain how music induces emotions in listeners.

More recently, Trost and colleagues [4] presented an in-depth discussion of existing theories explaining the central role of rhythmic entrainment in emotion elicitation in humans. Several studies cited in their paper consider rhythmic entrainment as one of the components of music inducing emotions in the listeners. The authors also formulated some hypotheses on the link between perceptual, physiological, motor and social entrainment and the arousal of emotional states in both listeners and players. Moreover, their literature review shows that motor entrainment, the one we address in this paper, "is experienced as a desirable and pleasant state, which would explain our attraction toward activities that require entrainment or why we choose to organize some everyday activities and behavior in a rhythmical way" [4].

Several previous works, e.g., [12], [14], [15], [16], observed that music joint performance is characterized by (intentional and unintentional) systematic deviations from strict synchrony between parts played by different individuals. Furthermore, the fact that deviations from a strict synchrony are necessary for a successful rhythmic joint action was formalized in the Yoshida and colleagues' theoretical model [17]. They argued that musicians do not continuously maintain a perfectly locked synchrony during music playing. In particular, they claimed that, in real music performance, players cannot completely reach and maintain the maximum level of synchrony. Instead, an alternation of high and low levels of synchrony (that they called *soft entrainment*) is deemed to make a performance more vivid and impressive because such an unconscious process can

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2019.2951368, IEEE Transactions on Affective Computing

trigger emotions in the listener. Synchrony periodically increases and decreases, raising in correspondence of the most significant moments of the performance (usually at the beginning and at the end of musical phrases). This is in line with the claims of Trost and colleagues [4], confirming that music pleasantness, enjoyability, and positive feelings while playing or listening to music are induced not only by maintaining a perfect motor synchrony, but by continuously reaching (and then, of course, slightly loosing) it.

2.3 Measuring interaction in music joint action

Most of existing research investigating entrainment in musical joint actions ground on interviews to group's members and manual annotations of audio-video recordings. Performances of duos, quartets, quintets (e.g., [16], [24], [25], [26], [27], [28]) as well as orchestras (e.g., [29], [30]) were addressed. More recently, a few studies attempted to quantify the dynamics of the interaction occurring in musical joint action. For example, Glowinski and colleagues [31] investigated the traits of the leadership in a string quartet. Results demonstrated that the leading musician is the one exhibiting the more regular and simple movements. Moreover, when a change occurs and the usual playing conditions of a music ensemble are artificially manipulated (for example, by switching the musicians' score or by asking them to play in an over-expressive fashion) the complexity of movement increases for all the musicians in the ensemble. Chang and colleagues [32] carried out a similar study to investigate interpersonal coordination beyond the note-to-note level by manipulating leadership in two quartets and the communication channels exploited by the musicians to produce music. Time series of body sway were collected through a motion capture system. They were analyzed by applying Granger causality to identify the magnitude and the direction of the information flow exchanged during the performances. Results showed that the prefixed leaders exerted a greater influence on the other players both when the information flow could be exchanged through the acoustic and the visual channels and, at a minor extent, when the visual channel was inhibited. Moreover, the authors found a significant positive correlation between the overall degree of body sway coupling (i.e., the average Granger causality of overall interpersonal body sway coupling) and the self-assessment of the goodness of the music performance.

Regarding larger groups, the focus was mainly on the mechanisms musicians exploit to coordinate among them or with their leader. Himberg and Thompson [33] studied choir performances to understand how singers can synchronize their movements and how expert singers influence novice singers' movements. They recorded singers' head and feet movements and they carried out a cross-correlation analysis revealing that tight synchrony occurs more among expert than novice singers, and that, due to the mutual adaptation of the singers, synchrony continuously evolves during performance. Furthermore, they also investigated cultural differences in bodily expression of entrainment, especially looking at corporeal representations of beat and metre [34]. Luck and Toiviainen [35] examined conductor's gesture features in an ecological setting. Their goal was to find out which features ensemble musicians synchronize their performance with. The conductor's gestures were recorded and 12 features (3D positions, 3D velocity and acceleration components, speed, magnitude of acceleration, and magnitude of acceleration along the movement trajectory) were extracted and cross-correlated with the beat of the ensemble's performance. The authors argued that the ensemble's performance tended to be most highly synchronized with periods of maximal deceleration along the trajectory, followed by periods of high vertical velocity. They also investigated the relationship between the kinematics of a conductor's expressive gestures and the ratings of perceived expression [36]. For this purpose, a point-light display representation of 2 conductors was presented to 24 participants who were asked to provide continuous ratings of valence, activity, power, and overall expression of gestures. Results showed that higher levels of expressivity were conveyed by gestures having increased amplitude, greater variance, and higher speed movement. D'Ausilio and colleagues [37] analyzed orchestra performances of some Mozart pieces. They recorded violinists' and conductors' movement kinematics and they searched for causal relationships among musicians by using Granger causality. Their results showed that an increase of conductor-to-musicians influence, together with a reduction of musician-to-musician coordination, is associated with an increased quality of execution, as assessed by music experts' judgments.

With respect to the related work, our research addresses a music ensemble large enough to need the presence of a conductor. It does not directly measure, however, the movement of the conductor and its kinematics as in [35]. Rather, it starts from the kinematics of the movements of the players for computing measures that can characterize entrainment and its variation over time. Similarly to [36], but differently from [33], our study takes also into account the affective dynamics of a musical performance. Differently from [36], however, our analysis does not focus on the expressive gestures the conductor performs, but rather it investigates how such affective dynamics is reflected into variations of entrainment and whether changing the conductor has an effect on such variations. Finally, similarly to the work of D'Ausilio and colleagues [37], our study includes external ratings of the music performance provided by music experts. Our work, however, does not focus on the causal relationships that establish between musicians and between conductor and musicians. Rather, it investigates and quantifies entrainment and its variation in the group, i.e., a primary variable such causal relationships operate on.

3 KURAMOTO MODELS

3.1 The original Kuramoto model

The original Kuramoto model is a mathematical model broadly used to study synchrony in large populations of coupled oscillators. Given a population of N oscillators, having natural frequencies ω_i exhibiting a probability density $g(\omega)$, the dynamics of each oscillator is generally governed by:

$$\dot{\theta}_i(t) = \omega_i + \sum_{j=1}^{N} C_{ij} sin(\theta_j(t) - \theta_i(t)) \quad i = 1, ...N, \qquad (1)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2019.2951368, IEEE Transactions on Affective Computing

where $\dot{\theta}_i(t)$ is the rate of change of phase; C_{ij} and $(\theta_j(t) - \theta_i(t))$ are the coupling strength and the phase difference between the i^{th} and j^{th} oscillators, respectively; $sin(\theta)$ is the coupling function. On the one hand, each oscillator tends to run independently at its own frequency. On the other hand, the coupling strength tends to synchronize it to all the others. When the coupling is sufficiently weak, the oscillators run incoherently, whereas beyond a certain threshold collective synchrony emerges spontaneously [38]. Kuramoto ran his model under the following assumptions: a very large number of nearly identical oscillators; mean-field (i.e., *all-to-all* coupling among oscillators and $C_{ij} = C/N \ \forall i, j$); time-independent coupling; unimodality and symmetry of the probability density of the natural frequencies $g(\omega)$. Within these assumptions Equation 1 becomes:

$$\dot{\theta}_i(t) = \omega_i + \frac{C}{N}\sum_{j=1}^{N} sin(\theta_j(t) - \theta_i(t)) \quad i = 1, ...N, \tag{2}$$

This model was successfully exploited in different contexts, such as in neurology [39].

In the Kuramoto model, the collective dynamics of the whole population of oscillators is measured by the *complex order parameter* [38]:

$$re^{I\psi} = \frac{1}{N}\sum_{j=1}^{N} e^{I\theta_j} \tag{3}$$

where $I = \sqrt{-1}$, $0 \leq r \leq 1$ measures the phase coherence of the population, and ψ is the average phase. The complex order parameter is the collective rhythm produced by the whole population of oscillators [40]. If $r \approx 1$, all oscillators are phase locked and act as a single "giant" oscillator. If $r \approx 0$, they add incoherently and no common rhythm is produced. When two or more interacting populations are modeled, an *overall order parameter* can be computed by combining together the order parameters r of each single population. The overall order parameter R is defined as the time-average of the single order parameters r.

To overcome some issues with the order parameter R, Kiss and colleagues [39] conceived another parameter: the *Synchronization Index* (*SI*). SI is very sensitive to interpopulation synchronization and it is defined as follows (for two populations of oscillators):

$$SI = \frac{\gamma(\langle r_1 \rangle + \langle r_2 \rangle)}{2} \tag{4}$$

where: γ measures the phase synchrony between the two groups and is computed as $\gamma = 1 - \frac{S}{S_{max}}$, with S being the Shannon entropy of the cyclic phase distribution and S_{max} the maximal entropy [41]; γ has an approximately unitary value for phase-locked groups and a value close to zero for phase-drifting; r_1 and r_2 are the Kuramoto's complex order parameters computed on the two populations of oscillators.

3.2 The generalized Kuramoto model

In realistic situations such as, e.g., neuronal networks, some generalizations of the Kuramoto model were developed to overcome the limitations due to the quite strict hypotheses the classical model is grounded on. Generalizations mainly concern the coupling function, the network structure, and time-variation of model parameters. Modified order parameters were also conceived.

1) Coupling function
Kuramoto chose *sinus* as coupling function in the formulation of his model. Nevertheless, other more complex periodic coupling functions were proposed in the literature. The simplest approach consists in expressing the coupling function $h(\theta)$ as a Fourier series [42]:

$$h(\theta) = \sum_{k=1}^{\infty}(h_k^s sin(k\theta) + h_k^c cos(k\theta)) \tag{5}$$

where θ is the difference between phases as in Equation 2, and h_k^s and h_k^c the series coefficients.

2) Network structure
Real world networks of oscillators typically show a more complex structure than all-to-all coupling. To study such kinds of networks, each oscillator is represented as a node and a $N \times N$ adjacency matrix describing all the connections is incorporated in the model formulation [43], [44]. The coupling between nodes i and j in the network is thus expressed as $C_{ij} = a_{ij}\sigma_{ij}$, being σ_{ij} the coupling strength and $A = \{a_{ij}\}$ the adjacency matrix.

3) Time variation of parameters
This namely includes time-variation of the natural frequency of the oscillators (i.e., $\omega_i = \omega_i(t)$) and time-variation of the coupling strength (i.e., $C_{ij} = C_{ij}(t)$) (see e.g., [45]).

4) Modified order parameters
Theoretical studies on the generalizations of the Kuramoto model also proposed modified order parameters. For example, the generalization of the model expressed by Equation 5 is accompanied by the definition of an order function, which is a generalization of the Kuramoto order parameter [42]. Nevertheless, the literature includes studies where the classical Kuramoto order parameter is used, even when a generalized model is adopted. Restrepo and colleagues [46] proposed an order parameter that can be used when the population of oscillators is arranged in a network with a general connectivity (i.e., the assumption of all-to-all coupling is not required) and the coupling terms between pairs of oscillators may be different and negative, meaning that the coupling may even drive the oscillators to be out of phase. Such an order parameter is defined as:

$$r = \frac{\sum_{i=1}^{N} r_i}{\sum_{i=1}^{N} d_i} \qquad r_i e^{I\psi_i} = \frac{1}{N}\sum_{j=1}^{N} C_{ij}e^{I\theta_j} \qquad d_i = \sum_{j=1}^{N} C_{ij}$$

The interpretation of the Restrepo's order parameter is similar to the classical one (see [44]), i.e., $r \approx 0$ when the oscillators are not synchronized, $r \approx 1$ if the oscillators are fully synchronized.

3.3 Kuramoto model in music ensembles

Playing in an orchestra requires cooperation due to a shared goal, division of roles, and monitoring of progresses [10]. This clearly implies coordination of movement making a collective rhythm emerge. Each player has her own rhythm, but through the interaction with the other players belonging to her orchestra section or to other sections, or with the conductor, she adjusts her rhythm toward the collective one.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2019.2951368, IEEE Transactions on Affective Computing

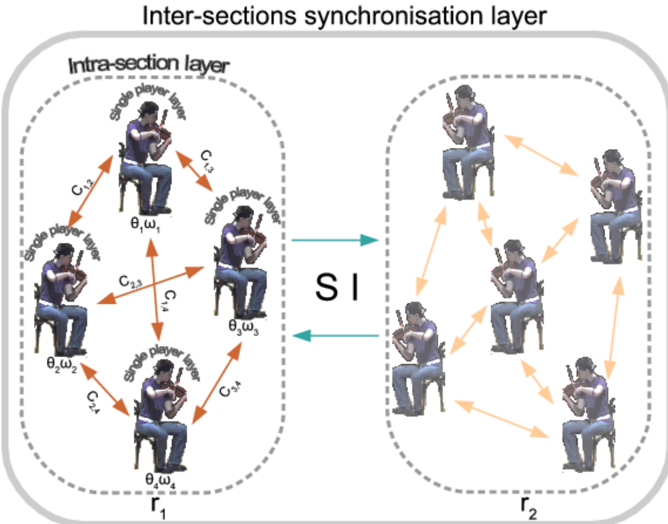


Fig. 1: How we apply the Kuramoto model to study entrainment in a musical joint action involving an orchestra playing a music piece. Two sections, one on the left and one on the right of the diagram, play together and interact. Each musician is modeled as an oscillator (Single player layer). Orange arrows represent the model of each section (Intra-section layer), where C_{ij} is the influence (coupling strength) of musician i on musician j. Blue arrows represent interactions between sections (Inter-sections layer), where the Synchronization Index (SI) is computed as a combination of the synchronizations (r_1 and r_2) within the single sections.

Figure 1 displays how we propose to apply the Kuramto model for studying entrainment in a musical joint action involving an orchestra playing a piece. Three layers take into account the dynamics of each single player (*Single player layer*), entrainment within a section (*Intra-section layer*), and entrainment between sections (*Inter-sections layer*).

At the Single player layer, each player is modeled as a single oscillator. The hypothesis to consider musicians as oscillators has been already successfully proposed in the literature: for example, Loehr and colleagues [47] proposed a temporally discrete oscillator model as a mathematical simplification retaining certain universal properties of neural oscillation including synchrony at multiple periodicities.

At the Intra-section layer, the coupling strength between players i and j in an orchestra section is assessed. Analysis of coupling and adjustment of rhythm among the players (i.e., coupled oscillator) is addressed by applying the Kuramoto model. The choice of the Kuramoto model to analyze coupling within a group is also supported in the literature (e.g., [48]). Although the players in a section can be deemed nearly identical, mean-field cannot be assumed because the coupling strength among the players is not identical. Moreover, the connectivity arrangement is not expected to be all-to-all, since some pairs of players may not be coupled. This prevents using the basic Kuramoto model described in Section 3 (Equation 2). Consequently, a generalized Kuramoto model enabling to express coupling in terms of an adjacency matrix (e.g., [43], [44]) needs to be adopted. In this framework, the coupling strength between players i and j is expressed as a matrix $C_{ij} = a_{ij}\sigma_{ij}$, being σ_{ij} the coupling strength and $A = \{a_{ij}\}$ an adjacency matrix describing the topology of the section, i.e., the network of players (oscillators). It is worth noticing that the adjacency matrix A may for example model constraints that the specific organization of music playing in an orchestra may impose to the topology of the network (e.g., musical instruments that have or have not to play together).

At the Inter-sections layer, entrainment between different sections of the orchestra is estimated. This is computed in terms of synchrony between different groups of players (i.e., between different populations of oscillators). Overall order parameters can be used for this purpose as well as the Synchronization Index.

4 EXPERIMENT

4.1 Set up and data-set

An orchestra (2 violin sections, 9 musicians in total) was observed in an ecological rehearsal scenario. The players performed 2 musical pieces they knew and routinely rehearsed: Mozart's K550-1 and K136-1. In the following, these pieces are referred as *P1* and *P2*, respectively. Each piece was repeated by the players 3 times (*Take1*, *Take2*, and *Take3*), both with their usual orchestra conductor (*known*) and with another professional conductor they never played with (*unknown*). The average duration of P1 over the takes of both conductors was 113s ($SD = 9.7s$), the average duration of P2 was 48.50s ($SD = 1.5s$). Each piece was segmented in *parts* according to changes in dynamics, attacks, and pauses: P1 counted 4 parts ($\langle p1 \rangle = 24.25$s, $SD_{p1} = 2.1$s; $\langle p2 \rangle = 25.00$s, $SD_{p2} = 1.9$s; $\langle p3 \rangle = 13.50$s, $SD_{p3} = 1.3$s; $\langle p4 \rangle = 52.25$s, $SD_{p4} = 4.3$s) whilst P2 counted 2 parts ($\langle p1 \rangle = 21.33$s, $SD_{p1} = 0.5$s; $\langle p2 \rangle = 27$s, $SD_{p2} = 1.1$s). Further, musical phrases in each of these parts were isolated: P1 includes 23 phrases, and P2 includes 11 phrases. Table 1 reports the correspondence between parts, phrases and bars for the two pieces.

Movement data was captured with a 3-camera Qualisys Motion Capture system recording (at 240 fps) the trajectory of reflecting markers attached to: (i) the tip of the bows; (ii) the conductor's baton; (iii) the conductor's left index finger. The data set analyzed in this study consisted of the data of the 9 musicians. It comprised 324 data segments: $4\ (parts) \times 9\ (musicians) \times 2\ (conductors) \times 3\ (takes)$ data segments for P1 and $2\ (parts) \times 9\ (musicians) \times 2\ (conductors) \times 3\ (takes)$ data segments for P2. To handle missing data due to markers occlusion, a spline method with continuous 3^{rd} order derivative was used for interpolation. Such interpolation was applied only when the percentage of missing data was less than 10% of the overall duration of the data segment and the missing data distribution showed far-between peaks (that is, missing data are not consecutive). If the percentage of missing data was higher than 10% the corresponding segment was discarded. Finally, a third order Savitzky-Golay filter was used to smooth trajectories. The data set for the analysis counted about 200 data segments (in particular, Take1 of P1 was discarded). The performance was also video and audio recorded. Figure 2 shows the orchestra playing during a rehearsal performance.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2019.2951368, IEEE Transactions on Affective Computing

TABLE 1: Correspondence between parts (columns 2 and 8), phrases (columns 3 and 9), and bars (columns 4-5 and 10-11) for the two classical music pieces P1 and P2 (Mozart's K550-1 and K136-1) exploited in the experiment. For each phrase, we specify the start and end bar (columns 4, 5, 10, and 11) and the type of content of each phrase (columns 6 and 12), 0 meaning that the phrase is characterized by single notes with pauses, 1 meaning that the phrase is characterized by notes without pauses.

Piece	Part	Phr.	Start bar	End bar	Type	Piece	Part	Phr.	Start bar	End bar	Type
P1	p1	1	1	5	1	P2	p1	1	1	4	0
P1	p1	2	6	9	1	P2	p1	2	5	8	1
P1	p1	3	10	11	1	P2	p1	3	9	12	1
P1	p1	4	12	16	0	P2	p1	4	13	16	1
P1	p1	5	17	20	0	P2	p1	5	17	24	1
P1	p2	6	21	24	1	P2	p2	6	25	29	0
P1	p2	7	25	27	1	P2	p2	7	30	33	0
P1	p2	8	28	37	0	P2	p2	8	34	37	1
P1	p2	9	38	42	1	P2	p2	9	38	41	1
P1	p3	10	43	47	0	P2	p2	10	42	46	1
P1	p3	11	48	51	0	P2	p2	11	47	59	1
P1	p3	12	52	54	1						
P1	p4	13	55	65	1						
P1	p4	14	66	70	1						
P1	p4	15	71	72	1						
P1	p4	16	73	75	1						
P1	p4	17	76	80	1						
P1	p4	18	81	84	1						
P1	p4	19	85	87	1						
P1	p4	20	88	90	1						
P1	p4	21	91	92	1						
P1	p4	22	93	94	1						
P1	p4	23	95	101	1						

Fig. 2: The orchestra during a rehearsal performance. Infrared reflecting markers are on the tip of the players' bows.

4.2 Quantitative analysis of performance

4.2.1 Single player layer

The z-component (i.e., the vertical coordinate) of the bow trajectory was selected to describe the dynamics of each player. The z-component was chosen for the analysis because this is the one exhibiting the largest variance and best displaying the oscillatory behavior of the player (which the x and y components may partially corrupt). The phase $\theta_i(t)$ of the z-component was computed by applying the analytic signal approach [49]. The natural frequency ω_i of each oscillator, i.e., of each player, was computed from the tempo (expressed in rad/sec) of the music piece the orchestra was playing and was considered as a constant along time. This approximation is reasonable as long as tempo does not change, i.e., the piece does not go through an *accelerando* or a *ritardando*. Whereas this is not usually the case in many music pieces, those selected for this study and the parts we analyzed have a quite steady tempo. In order to take into account inter-personal variability, the natural frequency of each player was randomly extracted from a Gaussian distribution centered on the music tempo and with unitary variance. The number N of oscillators (players) was low, but this did not affect the results as argued by theoretical studies and previous experiments (e.g., [45], [50]).

4.2.2 Intra-section layer

The generalized Kuramoto model described in [43], [44] was applied to compute the coupling strength C_{ij} between players i and j. We analyzed music pieces and parts where the music which is played does not change dramatically. Therefore, we approximated C_{ij} to be a constant along time. The model was fit to data by applying a regression through the origin, and goodness-of-fit was computed as referred in [51]. Goodness-of-fit was poor, that is, about 10%. To improve the model, a modified generalized Kuramoto model was applied, i.e., a more complex coupling periodic function was chosen: the first term of Fourier series with unitary coefficients, i.e., $h(\theta) = sin(\theta) + cos(\theta)$ (see Equation 5). Given these assumptions, the dynamics of the section was expressed by the following modified generalized Kuramoto model:

$$\dot{\theta}_i(t) = \omega_i + \sum_{j=1}^{N} C_{ij}[sin(\theta_j(t) - \theta_i(t)) + cos(\theta_j(t) - \theta_i(t))] \quad i = 1, \ldots N \quad (6)$$

A multiple regression through the origin was applied to two instances of Equation 6 – one for each violin section – for each part in each piece and in each take. It is noteworthy that C_{ij} cannot be estimated on a sample-by-sample basis because the number of equations is lower than the number of variables to be estimated. The regression coefficients C_{ij} that did not significantly predict $\dot{\theta}$ ($p > 0.05$) were set to zero, that is, the connection between player i and player j was pruned in the adjacency matrix $\{a_{ij}\}$. The resulting matrices of the regression coefficients C_{ij} were well-conditioned for every music piece, for every part, and for every section. Further, the ratio between the euclidean norm of residual and the output of the model was low for all the models. Therefore, the models fit the data. Goodness-of-fit was computed again for each model: the fitting of the models was satisfactorily high, between 34% and 88% (adjusted R^2). Table 2 shows the averaged adjusted R^2, conditioning numbers, and ratio between the euclidean norm of residual and the output of the models for each piece. Average was computed over the two sections and all the takes. All the values of the adjusted R^2 were significant at $p < 0.01$.

4.2.3 Inter-sections layer

Given the modified generalized Kuramoto model expressed by Equation 6, synchrony between the two orchestra sections can be computed using the order parameter defined by Restrepo and colleagues in [46]. In order to take into account possible negative couplings, the absolute value of the coupling strength $|C_{ij}|$ was used in the computation.

TABLE 2: The averaged adjusted R^2, conditioning numbers, and ratio between the euclidean norm of residual and the output of the models for each piece. Average was computed over the two sections and all the takes. All the values of adjusted R^2 were significant at $p < 0.01$.

Piece	Conductor	Avg R^2	Avg Cond. number	Avg. Ratio
P1	known	0.623 (SD=0.139)	4.73 (SD=2.18)	0.601 (SD=0.124)
P1	unknown	0.623 (SD=0.122)	4.62 (SD=1.73)	0.623 (SD=0.114)
P2	known	0.506 (SD=0.122)	3.27 (SD=0.566)	0.691 (SD=0.0966)
P2	unknown	0.532 (SD=0.106)	4.07 (SD=1.46)	0.677 (SD=0.0763)

The Synchronization Index SI (Equation 4) was then computed for each bar of each single musical phrase. To this aim, the time-average of the Restrepo's order parameters r_1 and r_2 for the two sections was computed on time windows corresponding to each bar each musical phrase is made of. Moreover, at the beginning of each phrase, except the first one, the time-average of the order parameters was computed on a time window centered on the starting sample of the phrase and having a width of one bar, i.e., half bar before the beginning of the phrase and half after it. This was done for taking into account possible strict synchrony deviation patterns around the beginning of the phrase. For the first phrase only the half bar after the beginning was considered. Similarly, at the end of each phrase, except the last one, the time-average of the order parameters was computed on a time window centered on the ending sample of the phrase and having a width of one bar. This was done for taking into account possible strict synchrony deviation patterns around the end of the phrase. For the last phrase only the half bar before the end was considered.

4.2.4 U-shaped or Monotonic Synchrony Pattern (UMSP)

In order to catch possible deviations from a strict synchrony between the orchestra sections, taking inspiration from [4], [14], [17], we looked at U-shaped or Monotonic Synchrony Patterns ($UMSP$). We consider that a $UMSP$ emerges if and only if one of the following conditions occur:

- i The SI value follows a raw U-shaped pattern (i.e., first it is monotonically decreasing, then monotonically increasing);
- ii The SI value monotonically increases/decreases.

Given any musical phrase $\phi \in P1|P2$, consisting of a sequence of N bars with synchronization values $SI_i, i \in 1, ..., N$, we computed the presence of $UMSP(\phi)$ as follows:

Algorithm 1: $Presence\_of\_UMSP(\phi)$

$\delta = \frac{SI_N - SI_1}{N}$
$Presence = true$
for $k \in \{2, ..., N-1\}$ **do**
  if $SI_k > (SI_1 + \delta * (N-1))$ **then**
    $Presence = false$
return $Presence$

Figure 3 displays an example of $UMSP$ presence detection for music piece P1 (Mozart's K136-1) take 2, performed by the orchestra with the unknown conductor. In the example, the phrases exhibiting a $UMSP$ are highlighted in green, while those not exhibiting it are highlighted in red.

4.3 Qualitative assessment of the performances

36 external raters were recruited via email advertisements at academies of music in Italy. They were professional players

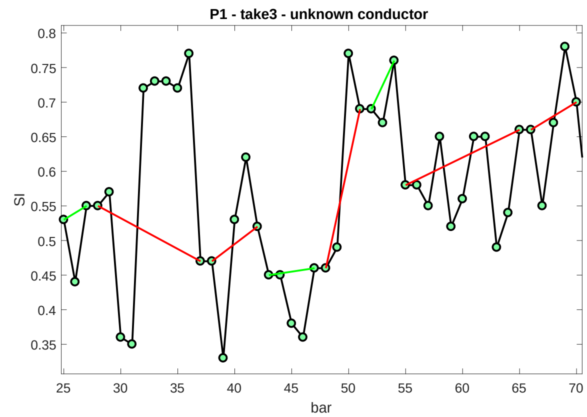


Fig. 3: Computation of the presence of $UMSP$s based on the variation of SI along the bars of part of piece P1, third take, performed by the orchestra with the unknown conductor. Phrases correspond to multiple bars, depending on the music content, as reported in Table 1. Green segments highlight phrases in which a $UMSP$ is detected (bars 25-27, 43-47, and 52-54), whilst red segments refer to phrases where it is not detected (bars 28-37, 38-42, 48-51, 55-65, and 66-70).

with more than 10 years of experience, composers, and music teachers. They were asked to listen to and to rate the orchestra performances from both a technical and an expressive point of view. Ratings were provided on two 5-step Likert items ranging from 0 (*poor*) to 4 (*excellent*). More specifically, the items were the following:

1. In terms of technical execution, how would you rate the performance? (referring to ability to play in time and to handle the technical difficulty of the piece)
2. In term of expressive execution, how would you rate the performance?

These items were conceived in the framework of a psychological study on the audience's experience of a music performance [52]. The raters provided their answer through an ad-hoc web application. At the beginning of the evaluation, raters were presented with a web page introducing the motivations of the study. They were also provided with written instructions on how to perform their task. Then, they were asked to listen to each audio excerpt and to rate it. They were asked to rate 10 excerpts each. An excerpt corresponded to one take of one of the two music pieces played by the orchestra with the known/unknown conductor, respectively. Raters could listen to each audio excerpt as many times as they wished before choosing their rating.

A Latin-square randomization procedure was adopted to present the stimuli.

Inter-raters reliability for each conductor was assessed by computing a 2-way mixed, consistency, average-measures Intra-class Correlation Coefficient (ICC) [53]. The resulting $ICCs$ were both in the good range [54] ($ICC_{known} = 0.65$, $ICC_{unknown} = 0.65$) showing that the participants similarly rated the audio excerpts for each conductor. Therefore, the average rating can be used in the analysis.

5 ANALYSIS AND RESULTS

Analysis was conducted as follows: (i) testing whether an effect of the conductor can be detected on the computed measures, (ii) exploring the possible relationship between the computed measures and the external ratings of the orchestra performance through linear models, and (iii) assessing whether an effect of the conductor can be detected on the model that better represented the relationship.

To take into account the role music structure may have on the joint action, analysis was conducted:

- On all the phrases of the music pieces (n=158);
- On the phrases at the beginning and at the end of each part each piece was divided into, i.e., the first phrase and the last phrase of each part (n=20);
- On the phrases at the middle of the pieces only, i.e., by excluding the first phrase and the last phrase of each part (n=138);
- On the phrases mainly characterized by single notes with pauses (n=38);
- On the phrases mainly characterized by notes without pauses (n=120).

5.1 Effect of the conductor on the computed measures

The effect of changing conductor on entrainment was tested by conducting paired t-tests on the SI values. A significant effect was only detected in the phrases characterized by single notes with pauses for the known (M=0.59) and unknown (M=0.49) conductor conditions; t(18)=2.68, $p < 0.05$ (see also Figure 4). Fisher's exact tests revealed that there was no the effect of the conductor on $UMSP$.

5.2 Analysis of the relationship between computed measures and external ratings

For checking for a possible relationship between the computed measures of SI and $UMSP$ and the external ratings, we applied several linear regression models and tested different hypotheses:

1) Relationship between SI and scores for items 1 and 2

Two simple linear regressions were carried out to investigate the relationship between SI and the scores attributed to item 1 and item 2, respectively. About item 1, a significant relationship was detected for all phrases, for central phrases, and for phrases characterized by single notes with pauses (see Table 3). The adjusted R^2 values show, however, that only a small amount of the variation in the attributed score can be explained by the model containing only SI. Concerning item 2, a significant relationship was detected for all phrases, for central phrases, and for phrases characterized by single notes with pauses (see Table 3). The adjusted R^2 values show, however, that only a small amount of the variation in the attributed score can be explained by the model containing only SI. Given that a relationship between SI and the scores attributed to both items was detected (even if the explained variance is indeed low), we further investigated whether and how the presence of $UMSP$s affects the parameters of the linear dependency models (intercept and angular coefficient).

TABLE 3: R^2 values for the simple linear regressions between SI and the scores attributed to item 1 and item 2. Statistical significance is reported via * (* $p < 0.05$, ** $p < 0.01$).

Phrases	R^2 item 1	R^2 item 2
All	0.05 (**)	0.03 (**)
External	n.s.	n.s
Central	0.08 (**)	0.06 (**)
With pauses	0.25 (**)	0.18 (**)
Without pauses	n.s.	n.s

TABLE 4: R^2 values for the multiple linear regressions through the origin between SI, $UMSP$, and the scores attributed to item 1 and item 2. Statistical significance is reported via * (* $p < 0.05$, ** $p < 0.01$).

Phrases	R^2 item 1	R^2 item 2
All	0.93 (**)	0.94 (**)
External	0.93 (**)	0.93 (**)
Central	0.94 (**)	0.94 (**)
With pauses	0.94 (**)	0.94 (**)
Without pauses	0.93 (**)	0.93 (**)

TABLE 5: R^2 values for the multiple linear regressions through the origin between SI, $UMSP$, the mutual interaction between SI and $UMSP$, and the scores attributed to item 1 and item 2. Statistical significance is reported via * (* $p < 0.05$, ** $p < 0.01$).

Phrases	R^2 item 1	R^2 item 2
All	0.96 (**)	0.96 (**)
External	0.94 (**)	0.94 (**)
Central	0.96 (**)	0.96 (**)
With pauses	0.96 (**)	0.96 (**)
Without pauses	0.96 (**)	0.96 (**)

2) Relationship between SI and scores attributed to items 1 and 2: effect of $UMSP$ on the presence of an intercept different from zero

Two multiple linear regressions through the origin were carried out to investigate whether the presence of $UMSP$s affects the presence of an intercept different from zero in the linear dependency between SI and the scores attributed to items 1 and 2, respectively. Independent variables were SI and $UMSP$ (dichotomous). A significant relationship was detected for both item 1 and item 2 for all phrases, for central phrases, for external phrases, for phrases characterized by single notes with pauses, and for phrases characterized by single notes without pauses (see Table 4).

3) Relationship between SI and scores attributed to items 1 and 2: effect of $UMSP$ on the presence of an intercept different from zero and on the angular coefficient

Two multiple linear regressions through the origin were

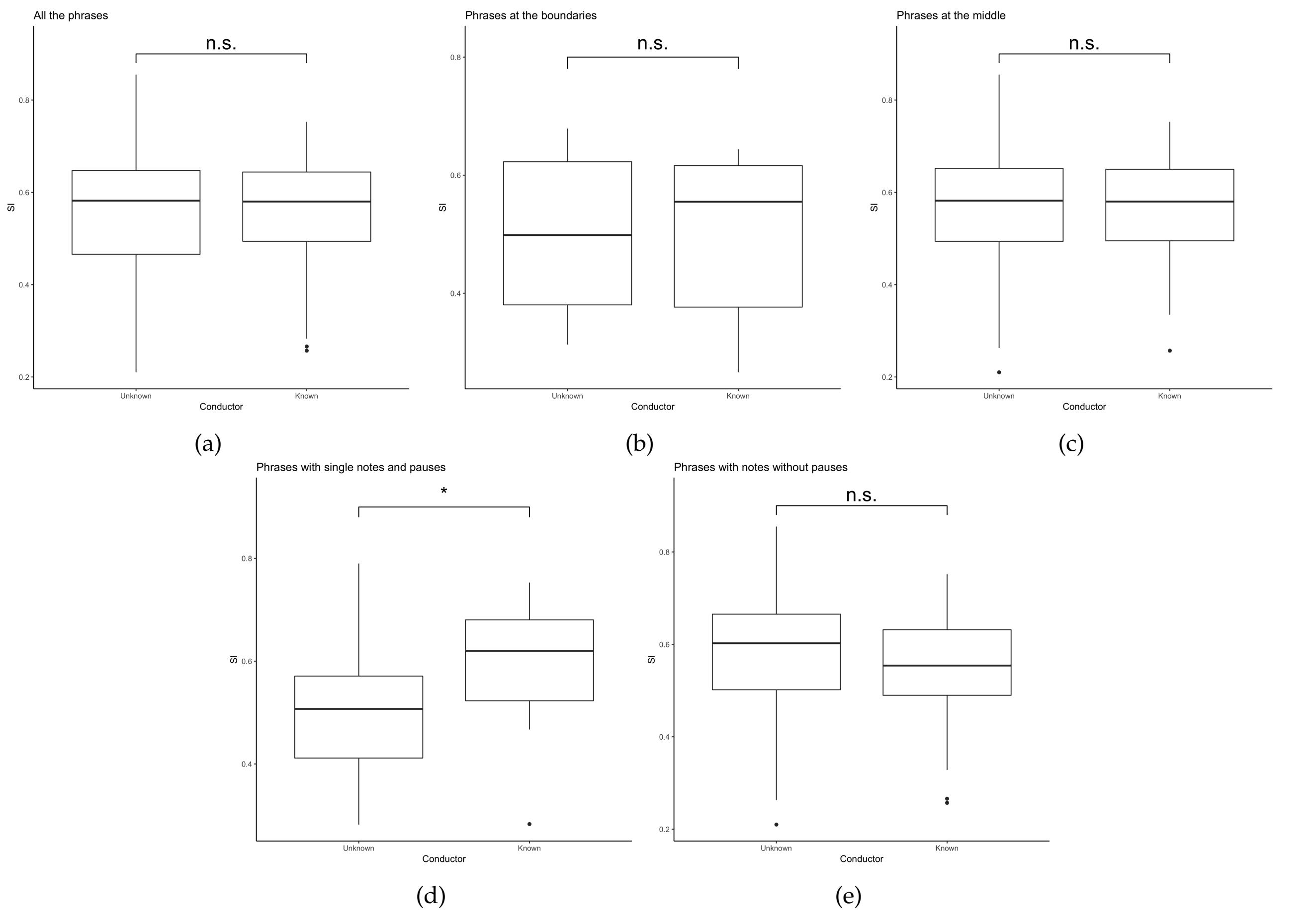


Fig. 4: Boxplots of the SI values computed for the known and for the unknown conductor for (a) all the phrases, (b) the phrases at the beginning and at the end of each part of each piece, (c) the central phrases of the pieces, (d) the phrases mainly characterized by single notes with pauses, and (e) the phrases mainly characterized by notes without pauses. Outliers are marked as full black dots. Statistical significance of the differences is reported via * (* $p < 0.05$, ** $p < 0.01$).

carried out to investigate whether the presence of $UMSP$s affects both the presence of an intercept different from zero and the angular coefficient of the linear dependency between SI and the scores attributed to items 1 and 2, respectively. Independent variables were SI, $UMSP$ (dichotomous), and the mutual interaction between SI and $UMSP$. A significant relationship was detected for both item 1 and item 2 for all phrases, for central phrases, for external phrases, for phrases characterized by single notes with pauses, and for phrases characterized by single notes without pauses (see Table 5).

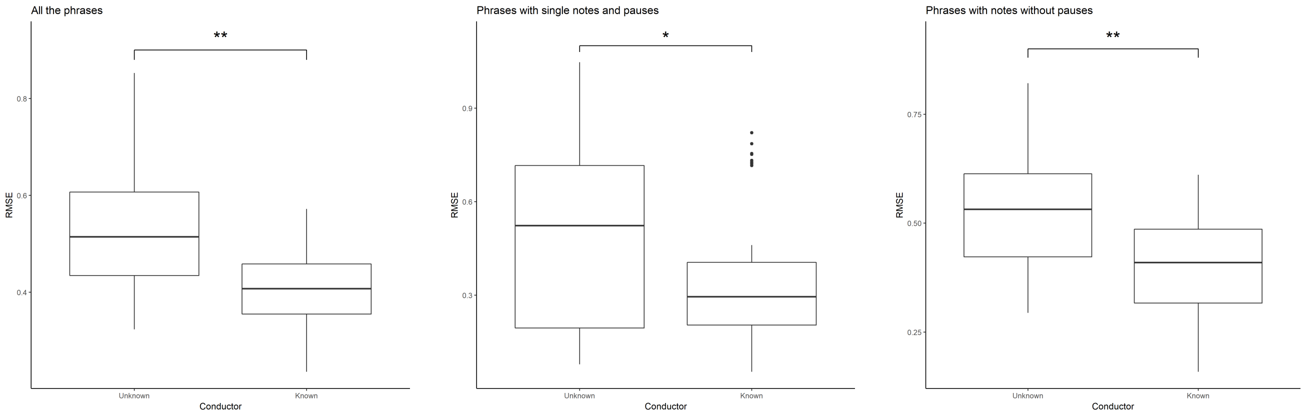
Results in 3) mean that the presence or absence of $UMSP$s determines two different linear dependencies between SI and the scores attributed to items 1 and 2. When a $UMSP$ is detected, the linear relationship between SI and the attributed scores passes through the origin and is characterized by a positive slope (slope coefficient = 4.101 for item 1 and slope coefficient = 3.851 for item 2), i.e., the attributed scores increase while SI increases. When a $UMSP$ is not detected, the linear relationship between SI and the attributed scores has an intercept different from zero (intercept = 2.810 for item 1 and intercept = 2.585 for item 2) and is characterized by a negative slope (slope coefficient = -0.624 for item 1 and slope coefficient = -0.542 for item 2), i.e., the attributed scores decrease from an initial value while SI increases. This model explains 96.0% of the variation in the scores attributed to both items. We therefore retained it for further assessing a possible effect of the change of conductor. For all the regressions, no evident nonlinear trends were detected in the diagnostic plots.

5.3 Effect of the conductor on predictability of the external ratings

For detecting a possible effect of the change of conductor, we analyzed whether the linear model in 3) performs differently for the two conductors. That is, we determined whether the conductor has an effect on the predictability of the attributed scores, being SI and $UMSP$ the predictors.

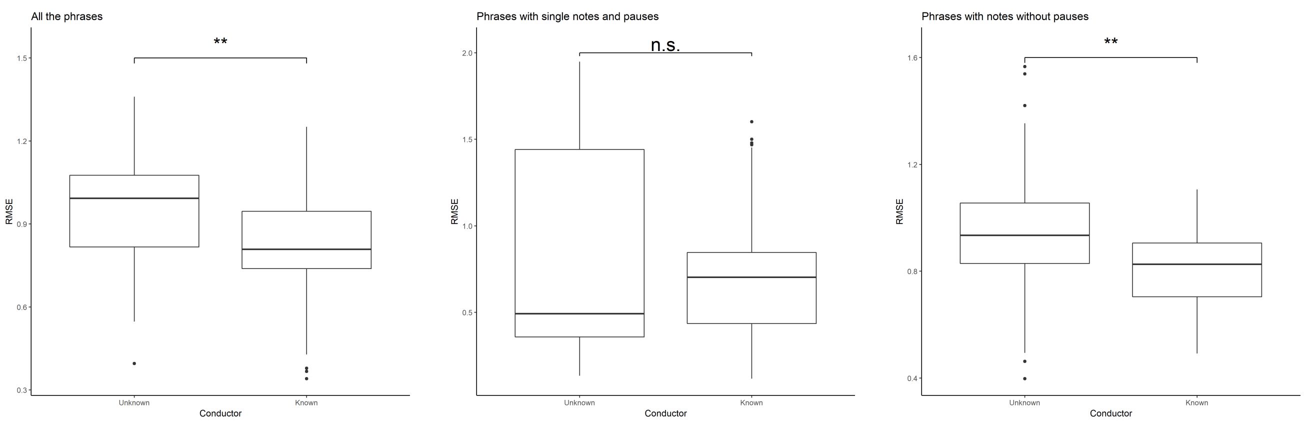
Datasets included data for all the phrases of the music pieces, for the phrases characterized by single notes with pauses, and for those ones characterized by notes without pauses. The size of the dataset including the phrases at the beginning and at the end of each part was indeed too small (n=20). So the distinction between peripheral and central phrases was not considered in this analysis. Predictability was assessed via a multiple-runs k-folds stratified cross-validation. We adopted 5 runs and 5 folds. Stratified random sampling was applied so that training and test datasets

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2019.2951368, IEEE Transactions on Affective Computing



(a) (b) (c)

Fig. 5: Boxplots of the RMSE values of the linear models predicting the scores attributed to item 1 depending on SI, $UMSP$, and the mutual interaction between SI and $UMSP$, for the known and for the unknown conductor. Boxplots refer to (a) all the phrases of the music pieces, (b) the phrases mainly characterized by single notes with pauses, and (c) the phrases mainly characterized by notes without pauses. Outliers are marked as full black dots. Statistical significance of the differences is reported via * (* $p < 0.05$, ** $p < 0.01$).



(a) (b) (c)

Fig. 6: Boxplots of the RMSE values of the linear models predicting the aggregated scores depending on SI, $UMSP$, and the mutual interaction between SI and $UMSP$, for the known and for the unknown conductor. Boxplots refer to (a) all the phrases of the music pieces, (b) the phrases mainly characterized by single notes with pauses, and (c) the phrases mainly characterized by notes without pauses. Outliers are marked as full black dots. Statistical significance of the differences is reported via * (* $p < 0.05$, ** $p < 0.01$).

contained an equal proportion of data referred to orchestra performances conducted by the two conductors. 25 regression models were computed on the training datasets and tested on the test datasets. For each dataset and for each model, Root Mean Square Error (RMSE) was computed separately on the samples of the test set referring to the known conductor and on the samples of the test set referring to the unknown conductor.

Item 1. Adjusted R^2 was significant for each model ($p < 0.01$). The coefficients of the models were significant as well ($p < 0.01$). No evident nonlinear trends were detected in the diagnostic plots. The average adjusted R^2 values for each dataset was 0.962 (all phrases), 0.960 (phrases with single notes with pauses), and 0.962 (phrases with notes without pauses). Paired t-tests were conducted on the RMSE values (see Figure 5). Results were significant for every dataset (all phrases: known (M=0.40), unknown (M=0.52), t(49)= -5.99, $p < 0.01$; phrases with single notes with pauses: known (M=0.36), unknown (M=0.46), t(49)=-2.10, $p < 0.05$; phrases with notes without pauses: known (M=0.41), unknown (M=0.53), t(49)=-6.53, $p < 0.01$;). The score attributed to item 1 for the performances conducted by the known conductor resulted overall more predictable than the score attributed to item 1 for the performances conducted by the unknown conductor.

Item 2. Adjusted R^2 was significant for each model ($p < 0.01$) as well as the coefficients of the models ($p < 0.01$). No evident nonlinear trends were detected in the diagnostic plots. The average adjusted R^2 values for each dataset was 0.964 (all phrases), 0.958 (phrases with single notes with pauses), and 0.965 (phrases with notes without pauses). Paired t-tests were conducted on the RMSE values. Results did not show any significant difference ($p > 0.05$).

Aggregated items. We finally repeated the same procedure on the aggregates scores, i.e., on the sum of the scores attributed to item 1 and 2. Adjusted R^2 was significant for each model ($p < 0.01$). The coefficients of the models were significant as well ($p < 0.01$). No evident nonlinear trends were detected in the diagnostic plots. The average adjusted R^2 values for each dataset was 0.964 (all phrases), 0.960 (phrases with single notes with pauses), and 0.964 (phrases with notes without pauses). Paired t-tests were conducted on the RMSE values (see Figure 6). Results were significant for all phrases (known (M=0.81), unknown (M=0.94), t(49)= -2.97, $p < 0.01$) and for the phrases with notes without pauses (known (M=0.79), unknown (M=0.94), t(49)=-2.89, $p < 0.01$). The aggregated score attributed to the performances conducted by the known conductor resulted more predictable than the aggregated score attributed to the performances conducted by the unknown conductor, but this does not hold for phrases with single notes with pauses (i.e., the paired t-test for those phrases was not significant).

6 DISCUSSION AND CONCLUSION

The analysis of the effect of the change of conductor on SI and $UMSP$ yields significant results only for phrases characterized by single notes with pauses. As literature on analysis of music ensembles points out, whereas macroscopic changes may have a clearly detectable effect (e.g., playing solo vs. playing in an ensemble, see for example, [55], [56]), important but less dramatical perturbations can also produce effects. These effects are however quite small, difficult to detect by means of technologies aiming at measuring behavioral variables, and requiring sophisticated analysis techniques (e.g., see [31], [57] for some examples concerning string quartets). The effect on SI found in phrases characterized by single notes with pauses shows that when synchrony needs to be more accurate – as in phrases including many isolated notes that have to be played exactly at the same time by all the players – the effect of change emerges and becomes measurable. The quite small difference in the mean values further supports these findings, i.e., the professional players did not dramatically deviate from the "expected" performance, but the requirements of some specific phrases made this subtle effect become more visible.

The analysis of the relationship between the computed measures and the scores raters attributed to each item revealed that the linear dependency of the scores on SI is twofold and depends on $UMSP$. Moreover, the slope coefficient changes its sign in the two cases. These findings show that, on the one hand, when entrainment is subject to variations, an increase of entrainment is reflected in an increase of the attributed scores. On the other hand, the absence of $UMSP$s results in the attributed score decreasing while SI increases. In other words, the more performers are entrained, but with fluctuations of entrainment, the higher is the score the performance gets. Conversely, the more performers are perfectly synchronized (without fluctuations of entrainment), the lower is the score the performance gets. This result is coherent with background theories about deviation from strict synchrony and reflects the common understanding that a too precise and mechanic performance is perceived as "cold" by external raters. Whilst not being an exhaustive argument in support of these theories, our findings seem to confirm this point of view.

The analysis of the effect of the change of conductor on predictability of the external ratings showed that the predictions made on data referring to the known conductor resulted more accurate. A possible reason for this is that knowing the conductor and her conducting style may help players focusing more tightly on their performance in terms of synchrony and its variation. With the unknown conductor, the adaptation mechanisms performers apply to carry out the joint action in the novel condition require them to take into account other variables that also influence the attributed score. This ultimately affects the effectiveness of our computed measures as predictors of the external ratings.

The work presented in this paper suffers of some limitations, which partly come from the difficulties of data collection and annotation. For example, the analysis only focused on musicians' positional data. Including further features in the analysis may provide a more comprehensive understanding of the phenomenon and its underlying mechanisms. Moreover, separated audio recordings of each single music instrument would allow to carry out a multimodal analysis: for example, it would be interesting to consider how musicians synchronize intonation and dynamics. It could be indeed possible to extract features that can be treated as oscillatory signals, e.g., the pulse of loudness of notes in a musical phrase, and oscillations of pitch intonation around the reference pitch, not just limited to the well-known *tremolo* and *vibrato*. Concerning the Kuramoto model, even if the selected coupling function proved suitable for fitting the model to data, different and more sophisticated coupling functions could be tested (e.g., including further terms of the Fourier series). Regarding the selected measures for our computational approach, it should be noticed that the theories about deviation from strict synchrony are not well established yet and still subject to investigation. More sophisticated measures will need to be developed, following theoretical advances in this area. Our analysis at the moment focused on linear regression. Nevertheless, a deeper investigation including the possible application of nonlinear models, would be worth pursuing. Other indicators of the goodness of the music performance could be used as dependent variables, including e.g., self-assessment reports provided by the players and the conductors. Last but not least, a more extended study involving more conductors and more orchestras would be needed to further support the results with a stronger evidence.

6.1 Conclusion

We investigated whether and how a perturbation in a group, i.e., the change of the conductor of a music ensemble, influences entrainment and consequently the orchestra effectiveness. To this aim, we adopted a computational approach to measure synchrony and its variation over time and we tested whether an effect of the change of conductor is detectable in the obtained measures. We studied the relationship between such measures and external ratings of the orchestra performance, and we assessed whether an effect of the conductor is detectable on the model that better

represents such a relationship. Results showed that whilst the change of conductor had a limited significant effect on entrainment (the effect was detectable in the phrases mainly characterized by single notes with pauses only), a significant effect was found when synchrony and deviation from strict synchrony were used as predictors of the external ratings.
The findings of this study may have just scratched the surface of a deep and subtle phenomenon, that the technologies and the measures we adopted are not able to capture in its full complexity yet. This nevertheless opens several encouraging perspectives for future work, both in terms of experimental activities that may, for example, involve ensembles of non professional players and/or more suitable tasks, and in terms of technological developments that should enable more sophisticated quantitative measures (e.g., behavioral and physiological) having a more fine-grained resolution.

Acknowledgments

We thank Leonardo Badino, Massimiliano Breviglieri, Corrado Canepa, Paolo Coletta, Alessandro D'Ausilio, and Cian Doherty for fruitful discussions. We also thank conductors Sera Tokay and Pietro Borgonovo for their precious help.

References

[1] J. Phillips-Silver and P. Keller, "Searching for roots of entrainment and joint action in early musical interactions," *Frontiers in Human Neuroscience*, vol. 6, p. 26, 2012.
[2] M. T. Brannick, R. M. Roach, and E. Salas, "Understanding team performance: A multimethod study," *Human Performance*, vol. 6, no. 4, pp. 287–308, 1993.
[3] N. Sebanz, H. Bekkering, and G. Knoblich, "Joint action: bodies and minds moving together," *Trends in cognitive sciences*, vol. 10, no. 2, pp. 70–76, 2006.
[4] W. Trost, C. Labbé, and D. Grandjean, "Rhythmic entrainment as a musical affect induction mechanism," *Neuropsychologia*, vol. 96, pp. 96–110, 2017.
[5] M. A. Marks, J. E. Mathieu, and S. J. Zaccaro, "A temporally based framework and taxonomy of team processes." *Academy of Management Review*, vol. 26, pp. 356–376, 2001.
[6] G. A. Okhuysen, "Structuring change: Familiarity and formal interventions in problem- solving groups." *Academy of Management Journal*, vol. 44, pp. 794–808, 2001.
[7] H. Arrow, J. E. McGrath, and J. L. Berdahl, *Small groups as complex systems.* Thousand Oaks, CA:Sage., 2000.
[8] D. G. Ancona and C. L. Chong, "Entrainment: Pace, cycle, and rhythm in organizational behavior," in *Research in organizational behavior*, B. M. Staw and L. L. Cummings, Eds. Greenwich, CT: JAI Press., 1996, ch. 18, pp. 251–284.
[9] S. E. Humphrey, F. P. Morgeson, and M. J. Mannor, "Developing a theory of the strategic core of teams: A role composition model of team performance," *Journal of Applied Psychology*, vol. 94, pp. 48–61, 2009.
[10] P. Keller, "Joint Action in Music Performance ," in *Enacting Intersubjectivity: A Cognitive and Social Perspective on the Study of Interactions*, F. Morganti, A. Carassa, and G. Riva, Eds. Amsterdam: IOS Press, 2008, pp. 205–221.
[11] M. Leman, *The Expressive Moment*. The MIT Press, 2016.
[12] P. Janata, S. T. Tomic, and J. M. Haberman, "Sensorimotor coupling in music and the psychology of the groove." *Journal of Experimental Psychology: General*, vol. 141, no. 1, p. 54, 2012.
[13] J. A. M. Correa, M. K. Abadi, N. Sebe, and I. Patras, "Amigos: a dataset for affect, personality and mood research on individuals and groups," *IEEE Transactions on Affective Computing*, 2018.
[14] P. E. Keller, G. Novembre, and M. J. Hove, "Rhythm in joint action: psychological and neurophysiological mechanisms for real-time interpersonal coordination," *Phil. Trans. R. Soc. B*, vol. 369, no. 1658, p. 20130394, 2014.
[15] W. Goebl and C. Palmer, "Synchronization of timing and motion among performing musicians," *Music Perception: An Interdisciplinary Journal*, vol. 26, no. 5, pp. 427–438, 2009.
[16] A. Geeves, D. J. McIlwain, and J. Sutton, "The performative pleasure of imprecision: a diachronic study of entrainment in music performance," *Frontiers in Human Neuroscience*, vol. 8, p. 863, 2014. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnhum.2014.00863
[17] T. Yoshida, S. Takeda, and S. Yamamoto, "The application of entrainment to musical ensembles," in *II International Conference on Music and Artificial Intelligence (ICMAI)*, Edinburgh, Scotland, 2002.
[18] G. Volpe, A. D'Ausilio, L. Badino, A. Camurri, and L. Fadiga, "Measuring social interaction in music ensembles," *Philosophical Transactions of the Royal Society B*, vol. 371, p. 20150377, 2016.
[19] M. Clayton, K. Jakubowski, and T. Eerola, "Interpersonal entrainment in indian instrumental music performance: Synchronization and movement coordination relate to tempo, dynamics, metrical and cadential structure," *Musicae Scientiae*, vol. 23, no. 3, pp. 304–331, 2019.
[20] M. Doffman, "Making it groove! entrainment, participation and discrepancy in the 'conversation' of a jazz trio," *Language & History*, vol. 52, no. 1, pp. 130–147, 2009.
[21] M. Leman, *Embodied Music Cognition and Mediation Technology*. The MIT Press, 2008.
[22] P. N. Juslin, "From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions," *Physics of Life Reviews*, vol. 10, no. 3, pp. 235 – 266, 2013.
[23] P. Juslin, S. Liljeström, D. Västfjäll, and L. Lundqvist, "How does music evoke emotions? exploring the underlying mechanisms," in *Handbook of music and emotion: Theory, research, applications*. Oxford University Press, 2010, pp. 605–642.
[24] V. M. Young and A. M. Colman, "Some psychological processes in string quartets," *Psychology of Music*, vol. 7, pp. 12–16, 1979.
[25] A. Williamon and J. W. Davidson, "Exploring co-performer communication," *Musicae Scientiae*, vol. 6, no. 1, pp. 53–72, 2002.
[26] E. C. Goodman, "Ensemble performance," in *Musical performance: A guide to understanding*, J. Rink, Ed. Cambridge: Cambridge University Press, 2002.
[27] L. Ford and J. W. Davidson, "An investigation of members' roles in wind quintets," *Psychology of Music*, vol. 31, no. 1, pp. 53–74, 2003.
[28] E. C. King, "The roles of student musicians in quartet rehearsals," *Psychology of Music*, vol. 34, no. 2, pp. 262–282, 2006.
[29] R. R. Faulkner, "Orchestra interaction: Some features of communication and authority in an artistic organization," *Sociological Quarterly*, vol. 14, no. 2, pp. 147–157, 1973.
[30] W. G. Bennis and B. Nanus, *Leaders: The strategies for taking charge*. Harper and Row, 1985.
[31] D. Glowinski, P. Coletta, G. Volpe, A. Camurri, C. Chiorri, and A. Schenone, "Multi-scale entropy analysis of dominance in social creative activities," in *Proceedings of the 18th ACM international conference on Multimedia*. ACM, 2010, pp. 1035–1038.
[32] A. Chang, S. Livingstone, D. Bosnyak, and L. Trainor, "Body sway reflects leadership in joint music performance," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 114, p. 21, 2017.
[33] T. Himberg and M. Thompson, "Group synchronization of coordinated movements in a cross-cultural choir workshop," in *7th Triennial Conference of the European Society for the Cognitive Sciences of Music, (ESCOM). Jyväskylä, Finland: University of Jyväskylä.*, 2009.
[34] H. Tommi and M. R. Thompson, "Learning and synchronising dance movements in south african songs–cross-cultural motion-capture study," *Dance Research*, vol. 29, no. supplement, pp. 305–328, 2011.
[35] G. Luck and P. Toiviainen, "Ensemble musicians' synchronisation with conductor's gestures: An automated features-extraction analysis," *Music Perception: An Interdisciplinary Journal*, vol. 24, no. 2, pp. 189–200, 2006.
[36] G. Luck, P. Toiviainen, and M. Thompson, "Perception of expression in conductors' gestures: A continuous response study," *Music Perception: An Interdisciplinary Journal*, vol. 28, no. 1, pp. 47–57, 2010.
[37] A. D'Ausilio, L. Badino, Y. Li, S. Tokay, L. Craighero, R. Canto, Y. Aloimonos, and L. Fadiga, "Leadership in orchestra emerges from the causal relationships of movement kinematics," *PLoS ONE*, vol. 7, no. 5, p. e35757, 05 2012.
[38] J. A. Acebrón, L. L. Bonilla, C. J. Pérez Vicente, F. Ritort, and R. Spigler, "The kuramoto model: A simple paradigm for synchro-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2019.2951368, IEEE Transactions on Affective Computing

nization phenomena," *Reviews of Modern Physics*, vol. 77, no. 1, pp. 137–185, 2005.
[39] I. Kiss, M. Quigg, S.-H. C. Chun, H. Kori, and J. Hudson, "Characterization of synchronization in interacting groups of oscillators: Application to seizures," *Biophysical Journal*, vol. 1, no. 94(3), pp. 1121–1130, 2008.
[40] S. Strogatz, "From kuramoto to crawford: exploring the onset of synchronization in populations of coupled oscillators," *Physica D: Nonlinear Phenomena*, vol. 143, no. 1-4, pp. 1–20, 2000.
[41] M. Rosenblum, A. Pikovsky, J. Kurths, C. Schäfer, and P. A. Tass, "Phase synchronization: from theory to data analysis," in *Handbook of biological physics*. Elsevier, 2001, vol. 4, pp. 279–321.
[42] H. Daido, "Onset of cooperative entrainment in limit-cycle oscillators with uniform all-to-all interactions: bifurcation of the order function," *Physica D: Nonlinear Phenomena*, vol. 91, no. 1-2, pp. 24 – 66, 1996.
[43] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, "Synchronization in complex networks," *Physics Reports*, vol. 469, no. 3, pp. 93–153, 2008.
[44] W.-X. Wang, L. Huang, Y.-C. Lai, and G. Chen, "Onset of synchronization in weighted scale-free networks," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 19, no. 1, p. 013134, 2009.
[45] D. Cumin and C. Unsworth, "Generalising the kuramoto model for the study of neuronal synchronisation in the brain," *Physica D: Nonlinear Phenomena*, vol. 226, no. 2, pp. 181 – 196, 2007.
[46] J. G. Restrepo, E. Ott, and B. R. Hunt, "Synchronization in large directed networks of coupled phase oscillators," *Chaos*, vol. 16, 2005.
[47] J. D. Loehr, E. W. Large, and C. Palmer, "Temporal coordination and adaptation to rate change in music performance." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 37, no. 4, p. 1292, 2011.
[48] F. Alderisio, G. Fiore, R. N. Salesse, B. G. Bardy, and M. di Bernardo, "Interaction patterns and individual dynamics shape the way we move in synchrony," *Scientific Reports*, vol. 7, 2017.
[49] D. Gabor, "Theory of communication. part 1: The analysis of information," *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering*, vol. 93, no. 26, pp. 429–441, 1946.
[50] T. Frank and M. Richardson, "On a test statistic for the kuramoto order parameter of synchronization: An illustration for group synchronization during rocking chairs," *Physica D: Nonlinear Phenomena*, vol. 239, no. 23-24, pp. 2084 –2092, 2010.
[51] J. G. Eisenhauer, "Regression through the origin," *Teaching Statistics*, vol. 25, 2003.
[52] C. Doherty, R. Cowie, C. Fyans, J. Jaimovich, B. Knapp, M. Ortiz-Perez, and P. Stapleton, "Development of a quality of experience questionnaire for use during live music performances," *Frontiers in Psychology*, vol. submitted, 2013.
[53] K. McGraw and S. Wong, "Forming inferences about some intraclass correlation coefficients," *Psychological Methods*, vol. 1, 1996.
[54] D. Cicchetti, "Guidelines, criteria and rules of thumb for evaluating normed and standardized assessment instruments in psychology," *Psychological Assessment*, vol. 6, 1994.
[55] P. Papiotis, M. Marchini, A. Perez-Carrillo, and E. Maestre, "Measuring ensemble interdependence in a string quartet through analysis of multidimensional performance data," *Frontiers in Psychology*, vol. 5, p. 963, 2014.
[56] D. Glowinski, M. Mancini, R. Cowie, A. Camurri, C. Chiorri, and C. Doherty, "The movements made by performers in a skilled quartet: a distinctive pattern, and the function that it serves," *Frontiers in Psychology*, vol. 4, p. 841, 2013.
[57] A. M. Wing, S. Endo, A. Bradbury, and D. Vorberg, "Optimal feedback correction in string quartet synchronization," *Journal of The Royal Society Interface*, vol. 11, no. 93, 2014.

Giovanna Varni received the M.Sc. degree in biomedical engineering and the Ph.D. in electronic, computer, and telecommunications engineering from University of Genoa, Italy, in 2005 and 2009, respectively. She is Maitre de Conferences at LTCI, Télécom Paris, Institut polytechnique de Paris, France. Her research interests are in the area of human machine interaction, especially on social signals analysis and expressive gesture.

Maurizio Mancini obtained his Ph.D. in Computer Science in 2008 at the University of Paris 8, France. Since 2001 he has carried out his research activity in the framework of several EU projects in FP5-7 and H2020. In 2008 he joined DIBRIS at the University of Genoa, Italy, as a postdoctoral researcher. His research activity focuses on the definition and implementation of models and algorithms for automated expressive movement analysis and synthesis in the field of HCI.

Luciano Fadiga M.D., University of Bologna, Ph.D. in Neuroscience, University of Parma. Senior Researcher at the University of Parma since 1992. He is actually Professor of Human Physiology at the University of Ferrara and Senior Researcher at the Italian Institute of Technology. He has a long experience in electrophysiology and neurophysiology in monkeys (single-neuron recordings) and humans (transcranial magnetic stimulation, study of spinal excitability, brain imaging, recording of single neurons in awake neurosurgery patients).

Antonio Camurri Ph.D. in Computer Engineering, is a full professor at DIBRIS, University of Genoa. Founder and scientific director of InfoMus Lab and of Casa Paganini (www.infomus.org). He is a coordinator and local project manager of several EU projects (FP5-FP7, H2020, Culture 2007, Cost Actions). His research interests include multimodal intelligent interfaces and interactive systems; sound and music computing; computational models of expressive gesture, emotion, and social signals; multimodal systems for theatre, music, dance, museums, health.

Gualtiero Volpe received the M.Sc. degree in computer engineering in 1999 and the Ph.D. in electronic and computer engineering in 2003 from the University of Genoa, Italy. Since 2014 he is an Associate Professor at DIBRIS, University of Genoa. His research interests include intelligent and affective human-machine interaction, social signal processing, sound and music computing, modeling and real-time analysis and synthesis of expressive content, and multimodal interactive systems.