PHILOSOPHICAL TRANSACTIONS B

rstb.royalsocietypublishing.org

Review

Cite this article: Volpe G, D'Ausilio A, Badino L, Camurri A, Fadiga L. 2016 Measuring social interaction in music ensembles. *Phil. Trans. R. Soc. B* **371**: 20150377. http://dx.doi.org/10.1098/rstb.2015.0377

Accepted: 5 February 2016

One contribution of 15 to a theme issue 'Attending to and neglecting people'.

Subject Areas:
behaviour, cognition

Keywords:
social interaction, music ensembles, mirror neurons, computational approaches, synchronization, leadership

Author for correspondence:
Gualtiero Volpe
e-mail: gualtiero.volpe@unige.it

Measuring social interaction in music ensembles

Gualtiero Volpe[1], Alessandro D'Ausilio[2], Leonardo Badino[2], Antonio Camurri[1] and Luciano Fadiga[2,3]

[1]Casa Paganini—InfoMus, DIBRIS, Università degli Studi di Genova, Viale Causa 13, Genova 16145, Italy
[2]Istituto Italiano di Tecnologia, CTNSC IIT@UniFe, Via Fossato di Mortara 17/19, Ferrara 44121, Italy
[3]Section of Human Physiology, University of Ferrara, Via Fossato di Mortara 17/19, Ferrara 44121, Italy

GV, 0000-0003-0760-4627

Music ensembles are an ideal test-bed for quantitative analysis of social interaction. Music is an inherently social activity, and music ensembles offer a broad variety of scenarios which are particularly suitable for investigation. Small ensembles, such as string quartets, are deemed a significant example of self-managed teams, where all musicians contribute equally to a task. In bigger ensembles, such as orchestras, the relationship between a leader (the conductor) and a group of followers (the musicians) clearly emerges. This paper presents an overview of recent research on social interaction in music ensembles with a particular focus on (i) studies from cognitive neuroscience; and (ii) studies adopting a computational approach for carrying out automatic quantitative analysis of ensemble music performances.

1. Introduction

Human interaction plays a central role in shaping cognition and brain organization. Nevertheless, cognitive neuroscience investigations only recently realized that the study of cognition should be carried out within social interaction [1]. Nowadays, we witness a renewed interest in the study of the brain during real-time truly interactive scenarios [2]. In neurocognitive research, however, the subject's behaviour is rarely tested in a realistic and salient social interaction owing to the necessity to maintain tight experimental control.

Recently, D'Ausilio *et al.* [3] proposed that music could offer a unique solution to balance a rigorous experimental approach and ecological testing of cognition and high brain functions. More interestingly, musicians and especially ensemble instrumentalists are experts in a form of social interaction characterized by real-time non-verbal communication. Ensemble musicians train for years in order to refine skills that allow them to accurately encode and decode subtle sensorimotor non-verbal messages with the main purpose of establishing and maintaining a shared coordinative goal. In group-level musical coordination, individuals might be conceptualized as processing units embedded within a complex system (i.e. the ensemble), engaged in a joint action, and sharing technical, aesthetic and emotional goals. Each participant may thus non-verbally transmit sensory information while, in parallel, decoding others' behaviours. The sensory information generated by the sender is based on body movements and is then transferred through the visual (e.g. body sway, head motion), auditory (e.g. instrument sounds), and somatosensory channels (e.g. floor vibrations). With information flowing, participants may rely on predictive models to cope with the real-time demands of interpersonal coordination.

Musicians in an ensemble are engaged in a joint action. This constitutes one of the most exciting frontiers of cognitive neuroscience [2,4]. Coordinated action might be conceived as the synchrony/complementarity between actions performed by two or more individuals [5]. Action coordination requires the exchange of information among individuals to facilitate understanding and prediction of others' motor intentions in what can be considered a sensorimotor conversation. Therefore, coordinated action is the modulation of ongoing

THE ROYAL SOCIETY PUBLISHING

motor plans according to sensorimotor information we read from other participants in the interaction and the surrounding context. For these reasons, music ensembles have been considered a valuable model to investigate the complex dynamics of sensorimotor communication during joint action [3]. In fact, music ensemble performance constitutes a special example of joint action that possesses critical advantages. Most notably, music possesses important constraints that allow better experimental control such as the fact that performers move in a complex but yet formalized manner, following musical conventions and, in most cases, a musical score. With respect to other experimental scenarios, such constraints are an ecological way to both reduce the number of experimental variables and overcome technical issues. For example, the fact that performers strictly follow a musical score allows for comparisons between different ensembles, something that would be more difficult in analysing, for instance, teamwork at a meeting, where discussion and work in different teams may follow significantly different paths. Moreover, music enables offer opportunities for analysing subtle aspects of non-verbal communication. For example, music often explicitly aims to convey emotion to an audience, which is not usually the case in, say, computer gaming. Finally, scenarios such as dance and sport usually encompass full-body movements that develop in quite a big space, something more difficult to capture accurately with optical motion capture systems.

This paper presents an overview of recent research on social interaction in music ensembles. After surveying ongoing research in neuroscience investigating music ensemble performance with reference to the mirror neuron mechanism, the focus will then move to two specific investigation scenarios: small ensembles, such as string quartets, and orchestras. String quartets are deemed a significant example of self-managed teams, where all musicians equally contribute to a task. In bigger ensembles, such as orchestras, the relationship between a leader (the conductor) and a group of followers (the musicians) clearly emerges. Whereas most works in the past were based on manual capturing, annotation, and analysis of data, interest recently has increased around computational approaches and techniques for automatic quantitative analysis. While reviewing recent research in both investigation scenarios, hints will be provided on the computational techniques used to carry out quantitative analysis.

2. Mirror networks and music

The discovery of trimodal (motor, visual, and auditory) mirror neurons in the monkey ventral premotor cortex [6] encouraged studies of the audiovisuomotor properties of the human mirror mechanism. This putative mechanism is thought to map the acoustic representation of actions into the motor plans necessary to produce those actions. For instance, action-related sounds were shown to activate the inferior frontal gyrus [7], following a clear somatotopic organization in the premotor cortex if the sound is executed by the hand or the mouth [8]. Interestingly, tool sounds activate brain areas overlapping with those recruited when participants pantomimed the manipulation of the same tools [9]. Finally, listening to non-verbal vocalizations automatically engages the preparation of responsive orofacial gestures [10]. These brain activities are located in the posterior aspect of the inferior frontal gyrus, which on the left hemisphere, corresponds to Broca's area. In parallel, other studies investigated the brain mechanisms associated with the processing of music (another instance of action-related sound) with a particular focus on the role of expertise. In fact, musicians are a particular example of sensorimotor expertise. Experts have been particularly useful in cognitive neuroscience, being an interesting model of over-learned sensorimotor associations [11,12]. Indeed, playing music is a multifaceted ability requiring years of practice to master, as it is based on the coordination of many different skills. Expert musicians are indeed an extraordinary population to study how specific training can shape somatosensory [13], motor [14], and auditory representations [15], as well as multimodal integration brain networks [16]. Musicians allowed the investigation of long-term structural [17] and short-term functional [18] plastic changes in the brain. In fact, sensorimotor integration in music could be the result of repeated co-occurrence of action and perceptual processes, with the additional advantage of being spread with different degrees of expertise among the population [19]. In fact, it has also been shown that specific postural and kinematic features associated with playing one instrument shape the activity of sensorimotor cortical areas during music listening. This result further suggests that motor training affects music perception [20]. Musicians, through years of practice, develop a strong functional association between a musical note, its visual representation, and the movement required to produce it [21,22]. As a consequence, stronger motor activity in expert pianists was found while they were listening to piano pieces [23]. Additionally, fMRI studies confirmed that a brain network including motor, premotor, and supplementary motor areas, the inferior parietal lobule, as well as the superior temporal gyrus, was shared between perception and production of a musical piece [24]. It was also shown that naive subjects, after a short musical training, had increased sensorimotor co-activation during passive listening of trained pieces [25,26], and passive listening of a trained piece induces a specific corticospinal facilitation after 30 minutes of practice [27]. Interestingly, rhythmic complexity or music syncopation is an important structural factor in embodied and affective responses to musical groove [28], and, indeed, high-groove music increasingly engages the motor system in musicians [29]. In addition, the observation of a mute piano fingering error induces a somatotopic time-locked corticospinal excitability modulation [30]. Moving to an interactive scenario, corticospinal excitability was modulated when pianists, performing one duet part, were led to believe that they were interacting with an invisible co-performer playing the complementary part [31]. Interference with the motor system (via transcranial magnetic stimulation) impaired interpersonal coordination only when the complementary part had been trained [32], and reduced the ability to predict the time course of the partner's actions during a musical turn-taking task [33]. Therefore, it is now clear that expert musicians exploit a mirror-like mechanism [34,35]: listening to musical excerpts or observing music-related movements evoke in them the same motor representations required for the actual production of those melodies.

In ensemble musical performance, a sensorimotor conversation requires that all participants are able to send and receive subtle messages in the form of visuomotor and audiomotor events. This encoding/decoding process might be conceived of as a complex and hierarchical input–output mapping, ranging from sensorimotor mapping to the highest level of human action organization [36,37]. In this context, the frontoparietal circuit, endowed with mirror-like properties, might be particularly important in non-verbal communication between individuals, especially when coordinated action is central. Action mirroring, however, does not facilitate coordinated action *per se*; in fact, coordination may often require the execution of different/complementary actions among participants. In agreement with this view, it has been demonstrated, however, that complementary action observation may recruit the human mirror neuron areas to a greater extent [38]. This result, in agreement with the behavioural requirements for sensorimotor interaction, suggests that the human mirror mechanism might be tuned for action coordination rather than for simple action mirroring. These results are in line with an interactive account of sensorimotor transformations forming the basis for human social behaviour. In fact, the mirror mechanism may provide the substrate to represent others' action in visuomotor coordinates rather than in symbolic codes. Such direct sensorimotor coding of others' action may result extremely effective in anticipating the motor consequences of others' actions and hence be used for fast online action inhibition, action selection, or action modulation, to best fit our own behaviour within interpersonal coordination. This whole process, based on the existence of a superordinate joint action goal, may be at the basis of non-verbal sensorimotor communication. According to us, for the case of orchestras and quartets, the investigation of sensorimotor non-verbal communication is fundamental to understand how to achieve the joint technical, aesthetic, and emotional goals of music.

3. Social interaction in small ensembles

In the framework of small music ensembles, many research efforts have been devoted to analysis of string quartets. The string quartet is the most significant chamber ensemble in Western music [39]. Moreover, Tal-Shmotkin & Gilboa [40,41] showed that the teamwork in a string quartet resembles those in various organizational units, with particular reference to self-managed teams. These are defined as a group of interdependent individuals, acting within an organizational setting, who self-regulate their behaviour in order to perform a relatively whole task assigned to them as a group [42]. Moreover, as Davidson & Good [43] note, from a musical point of view, string quartets are an interesting case study, because all musicians contribute similar elements to the musical performance and they use comparable instrumental techniques. Thus, cohesion and performance quality are more explicitly dependent on social dynamics in the group, rather than on technical aspects of playing music.

Several studies, mainly from psychology of music, investigated social interaction in string quartets. In their landmark study, Murnighan & Conlon [44] analysed 20 British string quartets using semi-structured interviews and observations. The authors found that successful quartets developed long-term collaborations, established effective non-verbal communication during rehearsals (e.g. by acting instead of talking to face a problem), were more democratic when managing internal conflict, reported more positive feelings before a concert, and were less worried about reviews. Besides performance issues, where the leadership of the first violin emerged, cohesion in the group depended also on merely organizational issues: the first violin was the leader in administrative issues too and the second violin was responsible for offering social support to the group. Davidson & Good [43] analysed social and musical coordination between members of a student string quartet in rehearsal and performance conditions. They found that social and musical coordination depends on many factors, including both socio-emotional issues (e.g. personal concerns, performance anxiety) and musical coordination issues related to content and process (e.g. the technical aspects of the music to be played and the mechanism adopted to achieve coordination such as the non-verbal gesture used to provide signals about timing). Other recent studies concerning string quartets from the point of view of psychology of music and specifically addressing social aspects include, for example, the work by King [45] focusing on collaboration between student musicians in a wind quartet, saxophone quartet, and string quartet and investigating the common roles assumed by student musicians, with particular reference to leadership. Seddon & Biasutti [46] revealed six modes of communication and two levels of attunement employed between members of a professional string quartet during rehearsal and performance and indicated that musicians were able to become empathetically attuned and produce spontaneous musical variations during performance. Schiavio & Høffding [47] presented a study based on qualitative interviews with the Danish string quartet and aiming at exploring the role of pre-reflective, embodied, and interactive intentionality in joint musical performance. Their analysis suggests that 'expert musicians' experience of collective music-making is rooted in the dynamical patterns of perception and action that co-constitute the sonic environment(s) in which they are embedded', i.e. that the cognitive processes involved are grounded in the concrete interactions of the players. String quartets were also analysed from a management and organizational point of view. For example, Butterworth [48] investigated the teamwork of the Detroit string quartet, using observations and interviews, and found that social–organizational factors (e.g. autonomy, norms of conduct, collective responsibility) were important elements in the work of the quartet and contributed to its success.

Psychological research on social interaction in ensemble performance is not limited to string quartets, but also addresses other ensembles. Keller & Appel [49] analysed individual differences in auditory imagery and temporal coordination in piano duos. This is the same music ensemble Goebl & Palmer [50] selected for investigating the influences of auditory feedback, musical role, and note ratio on synchronization. Analysing jazz ensemble playing, Berliner [51] studied the ability of 'finding the groove' as the ability to find a collective energy for the music. This can be viewed as a specific instance of the general 'mutual tuning-in relationship', i.e. the ability of a performer to anticipate the musical actions of the others, when such actions are available, e.g. performers are co-present [52]. In an analysis of a jazz quartet reported by Fuller [53] tuning-in depends on shared knowledge and rules which develop over time as long as each musician learns what to expect from the others and how to meet such expectations. Keller [54] developed a theoretical framework addressing ensemble performance in general and assuming that 'three core cognitive–motor skills determine the quality of real-time interpersonal coordination during music ensemble performance'. The first one relates to anticipatory mechanisms; the second one 'concerns the process of dividing attention between one's own actions and those of others while monitoring the overall, integrated ensemble output'; the third one grounds on the adaptive mechanisms allowing performers to react to variations. The framework was extended in [55], combining a descriptive level encompassing quantitative measures of behavioural cues and an explanatory level dealing with the psychological mechanisms that enable ensemble performers to achieve interpersonal coordination while simultaneously displaying the required flexibility.

Whereas most of such studies are based on manual annotation and analysis, interest has recently grown around computational approaches aimed at automatically computing features and using them to analyse different aspects of social interaction in small music ensembles. The string quartet is again a common target for analysis. For example, Moore & Chen [56] studied the interactive behaviour of two members of a skilled string quartet performing a selected musical passage that required both performers to play at a steady tempo and in synchrony. Angular velocity sensors were used to capture bowing movements. Results showed a high degree of synchrony. Moreover, interactive coupling between the players was found to be an essential component of joint performance. Employed techniques include signal-processing methods such as auto- and cross-correlation and convolution. Informally, auto- and cross-correlation are methods to look for similarity between observations within the same signal (autocorrelation) and across two different signals (cross-correlation), as a function of a time lag. Glowinski *et al.* [57] compared the expressive movement of the first violinists in two string quartets when playing in solo and ensemble conditions. Head movement patterns were quantified using a measure of entropy, and findings showed that head movements are more regular and predictable in ensemble playing than in solo playing. The selected entropy measure was sample entropy [58]. This is a nonlinear technique for measuring complexity, initially developed to quantify behaviour regularity and improved by Govindan *et al.* [59]. Wing *et al.* [60] asked two string quartets to repeatedly perform a short musical excerpt while introducing unrehearsed expressive deviations in terms of timing in their performance. They applied time-series analysis of successive tone onset asynchronies to estimate correction gains for all pairs of players. A first-order linear phase correction model [61,62] was adopted to quantify the correction gain each player had to perform to adjust her tempo and achieve synchrony. Results showed that on average, both quartets exhibited near-optimal gain. Contrasting patterns of adjustment between some pairs of players emerged, reflecting contrasting strategies of first-violin-led autocracy versus democracy. Another approach consists of measuring the degree to which individual musicians interact with and influence each other in a string quartet performance. Badino *et al.* [63] used granger causality [64] applied to body sway kinematics, to measure the leadership exerted by the first violin. Granger's method is an autoregressive technique used to establish causal relationships between time series—and thus applicable to the sensorimotor transfer of information between musicians. Leadership was modelled as the balance between how much more each musician was 'causing' movement of others than she was caused by the rest of the group. Interestingly, the introduction of an informational asymmetry, in the form of incompatible agogic notation provided to the first violin only, led to a significant reduction of his influence towards the rest of the musicians. The authors suggested that the introduction of extreme unexpected dynamical changes in a well-rehearsed piece is a very unusual scenario for professional musicians, leading to a disruption of the normal sensorimotor communication patterns. Papiotis *et al.* [65] analysed the audio and motion capture recordings (movement of the bow) of a quartet playing exercises under two experimental conditions: solo (i.e. each musician performs her part alone using a stripped-down version of the score) and ensemble. Features related to four different dimensions of the performance (intonation, dynamics, timbre, and tempo) were automatically extracted from the data. The level of interdependence between the musicians was estimated applying two linear and two nonlinear techniques to the time series of the features. Linear methods included Pearson product–moment correlation coefficient and granger causality. Nonlinear methods included mutual information—a measure of the mutual dependence between two variables quantifying the amount of information obtained about one variable through the other one [66,67]—and the nonlinear coupling coefficient for the estimation of directional couplings from time series [68,69]. Results showed that it is possible to automatically distinguish the two experimental conditions by quantifying interdependence between musicians in each of the performance dimensions studied. The nonlinear methods appeared to perform best. Moreover, by using the solo performance as reference, the authors could estimate the amount of interdependence established between the musicians across different exercises, and relate such a measure with the goal of the exercise.

Concerning examples of quantitative studies using computational approaches and addressing other small ensembles, the work on piano duos by Keller & Appel [49] mentioned above applied cross-correlation on kinematic data (velocity and acceleration profiles) retrieved from motion capture recordings. Varni *et al.* [70] applied recurrence quantification analysis to the head movements of violin duos in order to analyse synchronization and leadership. Recurrence quantification analysis is a method of nonlinear data analysis, quantifying the small-scale structures of recurrence plots, which present the number and duration of the recurrences of a dynamical system [71]. Findings showed that induction of a positive emotion in one player encompassed higher synchronization, whereas no clear leadership of one player emerged, confirming the hypothesis of egalitarian distribution of leadership in duos [72].

Some recent work faced analysis of small music ensembles in a music information retrieval (MIR; [73]) perspective, i.e. with the aim of extracting information such as musical style or genre. Whereas, on the one hand, such studies do not directly address social interaction issues; on the other hand, the audio features they develop and extract can also be applied to analysis of social interaction. For example, Abesser *et al.* [74] describe a collection of rhythm-related audio features used to estimate groove quality and interplay quality in ensembles made by electric guitar, bass guitar, and drums. Devaney *et al.* [75] developed a toolkit for accurately aligning monophonic audio to MIDI scores as well as extracting and analysing timing-, pitch-, and dynamics-related performance data from the aligned recordings. The toolkit was applied in studying intonation in three-parts singing.

4. Leader–follower relationship in orchestras

Orchestras are another particularly interesting instance of sensorimotor coordination. In fact, in contrast with quartets or other smaller ensembles, the social organization characterizing these groups is based on a clear and formalized leadership. The conductor is *de facto* the driving force behind the pattern of group information flow, at multiple levels. The conductor is not merely giving tempo to the players or the timing of attacks during public performances. In fact, the conductor is also selecting musicians, leading rehearsals, as well as deciding fine interpretation details of the pieces, based on personal taste and a philological erudition [76]. At the highest levels, the conductor is the major driver in fine-tuning all aspects of the performance [77]. The sensorimotor conversation between musicians and conductors is certainly important for musical performance. Following the traditional view that the conductor conveys mostly the expressive content of music, some studies have investigated the effectiveness by which conductors' gestures evoke the desired expressive content and the combination of kinematic features that maximize such classification [78]. Investigations of the temporal aspects of coordination have shown that musicians could effectively synchronize with point-light display of a conductor's gestures, provided that these comply with biological motion rules. Importantly, visual cues were as effective as auditory ones [79]. Furthermore, it was shown that conductors, even more than musicians, are capable of precise synchronization with an external visual stimulus. This fact suggests that orchestra conductors are specifically tuned to the detection of salient kinematic features of body gestures that allow temporal prediction and estimation [80]. Interestingly, one study investigated the perception of conductors' body movements under different viewing conditions. Video sequences presenting the conductors' faces induced higher expressiveness ratings than the arms-only conditions. At the same time, sequences showing the arms were judged higher in amount of information [81]. In agreement with this result, it has been shown that conductors' eye movements are directed at the score, but most of the score-reading was in anticipation of the music to be played. Interestingly, the longest anticipations were associated with segments of critical expressive importance for the piece [82]. Authors suggest that the conductor's eye movements were governed by two different timescales. A short one related to the coordination of instrumental performance and the other on a larger scale, and associated with her/his expressive musical conception of the piece. Therefore, orchestra music performance is a remarkable instance of social interaction in which the conductor uses her/his motor behaviour to drive the players towards a common aesthetic goal. Such interaction might happen at multiple timescales and potentially using whole-body movement features.

The collective behaviour of orchestras is a powerful model of interindividual non-linguistic communication among highly skilled individuals. Such a kind of non-verbal communication enables efficient transfer of information among musicians. In this context, D'Ausilio *et al.* [83] used violinists' and conductors' movement kinematics to search for causal relationships among musicians during real musical performances. Musicians were playing Mozart pieces they knew in advance, with two different conductors. The acceleration profiles of bow movements and conductors' right-wrist action were analysed with the Granger causality method [64]. Results showed that the increase of conductor-to-musicians causal influence together with the reduction of musician-to-musician coordination (an index of successful leadership) affect quality of execution, as assessed by musical experts' judgements. Rigorous quantification of non-verbal communication efficacy has always been complicated and affected by rather vague qualitative methodologies. The work by D'Ausilio *et al.* [83] demonstrated that the analysis of motor behaviours provides an effective tool to quantify the rather intangible concept of aesthetic quality of music, efficacy of non-verbal communication, and successful leadership in orchestral conducting. Gnecco *et al.* [84] showed how a reduced set of simple movement features—head movements—could be used to measure the levels of attention of the musicians with respect to the conductor and the music stand under various conditions.

5. Conclusion and open research challenges

As previous sections reported, much research effort has been devoted to shedding light on the mechanisms of social interaction in music ensembles from several perspectives (e.g. cognitive neuroscience, psychology of music, musicology, management and organizational sciences, information and communication technologies). Major results include a deeper knowledge of the mirror-like mechanisms involved in music performance, the development of quantitative approaches to the analysis of social interaction in music ensembles, e.g. with respect to features such as entrainment and leadership, and the availability of hardware and software platforms and tools for automatic data analysis and processing.

Many challenges remain open. Scientific issues include generalization of results, cross-cultural differences, and ecological validity. These entail several technical issues concerning data acquisition and analysis, e.g. with respect to recording protocols, availability of tools for synchronization and annotation of multimodal signals, and management of large datasets.

Generalization of results is possibly one of the most urgent challenges. Many factors may affect generalization. One of them is sample size. It is inherently difficult to recruit, record, and analyse many string quartets or many orchestras. Most studies thus focus on multiple performances by the same ensemble or by a few of them. The risk is to measure idiosyncratic features of that specific ensemble rather than general properties. Some studies involved a large number of ensembles (e.g. Murnighan & Conlon [44] studied 20 string quartets), but these studies are usually based on interviews and observations, rather than on data analysis of recorded performances. Another aspect concerns which movement is analysed and which movement features are used to describe it. Most studies focused on body motion kinematics only, and addressed the movement of just one single body part (e.g. the position of the upper end of the bow of a violin player). If, on the one hand, this approach has the clear advantage of granting good experimental conditions, on the other hand, such a specific focus may prevent generalization of results, because it misses the complexity of the ongoing coordinative behaviour of the ensemble. In other words, the conclusions remain somewhat limited by the context and the measures implemented in the experimental scenario.

Multi- and cross-cultural differences are also related to generalization of findings. Whereas most studies focused on Western music and on consolidated ensembles in the framework of Western music (e.g. the string quartet), a cross-cultural perspective should be adopted and some studies in this direction are available in the literature. For example, Clayton [85] studied the emergence, in an Indian music concert, of entrainment between individual parts that were explicitly intended to be uncoordinated, whereas Lucas *et al.* [86] demonstrated the emergence of entrainment between distinct groups of musicians and dancers in the Afro-Brazilian ritual Congado. An example of a computational approach to automated analysis of interpersonal entrainment in Indian music performance, applying cross-correlation measures, is available in [87].

Ecological validity poses significant organizational (e.g. availability of suitable locations such as a concert hall, recruitment of an audience) and technical challenges. A rigorous testing of interindividual coordination in an ecological scenario requires complex set-ups, needing long preparation time and making difficult the design of laboratory-like experiments with multiple repetitions of the same conditions as well as a large number of participants. For instance, see the pioneering works of Babiloni *et al.* [88,89] on how to record electroencephalographic signals from up to four musicians. Technical issues include, for example, the need to synchronize multimodal data streams accurately, and the extended time required for data post-processing and annotation.

Finally, especially with respect to ecological validity, one further research scenario is worth investigating which was not addressed in this paper: analysis of social interaction between musicians and audience. This scenario encompasses significant research questions such as how entrainment between musicians, expression of emotion, and co-creation affect audience experience.

The impact of the research surveyed in this paper is potentially broad, going far beyond the domain of music. Besides immediate application to management and organizational sciences, the development of computation models and techniques for automatic analysis of social interaction—nowadays, often referred to as social signal processing [90]—opens new research and application avenues. For example, such novel technologies can both support a more quantitative approach to research in psychology and musicology, and be embedded in future ICT applications for everyday life. Application domains include, e.g. embodied social media, user-centric media, music industry, education, cultural heritage, independent living, therapy, and rehabilitation. Above all, such novel technologies enable the design of social interfaces and computer systems capable of interacting socially. This represents a significant breakthrough with respect to the previous generation of computer systems and interfaces, which were usually designed for interaction with a single user and do not display any social capability. The availability of computer systems and interfaces endowed with social capabilities enables applications such as training and analysis of group dynamics (e.g. groups of employees or students) to improve teamwork, media content indexing, and development of active software agents, e.g. embodied companion agents.

Authors' contributions. G.V. and A.C. mainly contributed to the section on social interaction in small ensembles. A.D., L.B., and L.F. mainly contributed to the sections on mirror networks and music, and on the leader–follower relationship in orchestras. All authors contributed to the introduction and the conclusion. G.V. edited the whole paper.

Competing interests. We have no competing interests.

Funding. This work has been partially supported by EU Grants Siempre and Poeticon++.

Acknowledgements. We deeply thank Maestro Riccardo Muti for his inspiring collaboration, Rosario Canto for his help with data acquisition in orchestra scenario, and Corrado Canepa, Paolo Coletta, Donald Glowinski, Maurizio Mancini, and Giovanna Varni, who worked on analysis of social interaction in music ensembles in the EU-FET Siempre Project.

References

1. Schilbach L, Timmermans B, Reddy V, Costall A, Bente G, Schlicht T, Vogeley K. 2013 Toward a second-person neuroscience. *Behav. Brain Sci.* **36**, 393–414. (doi:10.1017/S0140525X12000660)
2. Hari R, Henriksson L, Malinen S, Parkkonen L. 2015 Centrality of social interaction in human brain function. *Neuron* **88**, 181–193. (doi:10.1016/j.neuron.2015.09.022)
3. D'Ausilio A, Novembre G, Fadiga L, Keller PE. 2015 What can music tell us about social interaction? *Trends Cogn. Sci.* **19**, 111–114. (doi:10.1016/j.tics.2015.01.005)
4. Sebanz N, Bekkering H, Knoblich G. 2006 Joint action: bodies and minds moving together. *Trends Cogn. Sci.* **10**, 70–76.
5. Sommerville JA, Decety J. 2006 Weaving the fabric of social interaction: articulating developmental psychology and cognitive neuroscience in the domain of motor cognition. *Psychon. Bull. Rev.* **13**, 179–200. (doi:10.3758/BF03193831)
6. Kohler E, Keysers C, Umiltà MA, Fogassi L, Gallese V, Rizzolatti G. 2002 Hearing sounds, understanding actions: action representation in mirror neurons. *Science* **297**, 846–848. (doi:10.1126/science.1070311)
7. Pizzamiglio L, Aprile T, Spitoni G, Pitzalis S, Bates E, D'Amico S, Di Russo F. 2005 Separate neural systems for processing action- or non-action-related sounds. *Neuroimage* **24**, 852–861. (doi:10.1016/j.neuroimage.2004.09.025)
8. Gazzola V, Aziz-Zadeh L, Keysers C. 2006 Empathy and the somatotopic auditory mirror system in humans. *Curr. Biol.* **16**, 1824–1829. (doi:10.1016/j.cub.2006.07.072)
9. Lewis JW, Brefczynski JA, Phinney RE, Janik JJ, DeYoe EA. 2005 Distinct cortical pathways for processing tool versus animal sounds. *J. Neurosci.* **25**, 5148–5158. (doi:10.1523/JNEUROSCI.0419-05.2005)
10. Warren JE, Sauter DA, Eisner F, Wiland J, Dresner MA, Wise RJ, Rosen S, Scott SK. 2006 Positive emotions preferentially engage an auditory–motor 'mirror' system. *J. Neurosci.* **26**, 13 067–13 075. (doi:10.1523/JNEUROSCI.3907-06.2006)
11. Calvo-Merino B, Glaser DE, Grèzes J, Passingham RE, Haggard P. 2005 Action observation and acquired motor skills: an FMRI study with expert dancers. *Cereb. Cortex* **15**, 1243–1249. (doi:10.1093/cercor/bhi007)
12. Aglioti SM, Cesari P, Romani M, Urgesi C. 2008 Action anticipation and motor resonance in elite basketball players. *Nat. Neurosci.* **11**, 1109–1116. (doi:10.1038/nn.2182)
13. Elbert T, Pantev C, Wienbruch C, Rockstroh B, Taub E. 1995 Increased cortical representation of the fingers of the left hand in string players. *Science* **270**, 305–307. (doi:10.1126/science.270.5234.305)
14. Pascual-Leone A, Nguyet D, Cohen LG, Brasil-Neto JP, Cammarota A, Hallett M. 1995 Modulation of muscle responses evoked by transcranial magnetic stimulation during the acquisition of new fine motor skills. *J. Neurophysiol.* **74**, 1037–1045.
15. Pantev C, Oostenveld R, Engelien A, Ross B, Roberts LE, Hoke M. 1998 Increased auditory cortical representation in musicians. *Nature* **392**, 811–814. (doi:10.1038/33918)
16. Stewart L, Henson R, Kampe K, Walsh V, Turner R, Frith U. 2003 Brain changes after learning to read and play music. *Neuroimage* **20**, 71–83. (doi:10.1016/S1053-8119(03)00248-9)

17. Schlaug G, Jäncke L, Huang Y, Steinmetz H. 1995 *In vivo* evidence of structural brain asymmetry in musicians. *Science* **267**, 699–701. (doi:10.1126/science.7839149)
18. Rosenkranz K, Williamon A, Rothwell JC. 2007 Motorcortical excitability and synaptic plasticity is enhanced in professional musicians. *J. Neurosci*. **27**, 5200–5206. (doi:10.1523/JNEUROSCI.0836-07.2007)
19. Münte TF, Altenmüller E, Jäncke L. 2002 The musician's brain as a model of neuroplasticity. *Nat. Rev. Neurosci*. **3**, 473–478.
20. Burunat I, Brattico E, Puoliväli T, Ristaniemi T, Sams M, Toiviainen P. 2015 Action in perception: prominent visuo-motor functional symmetry in musicians during music listening. *PLoS ONE* **10**, e0138238. (doi:10.1371/journal.pone.0138238)
21. Zatorre RJ, Chen JL, Penhune VB. 2007 When the brain plays music: auditory–motor interactions in music perception and production. *Nat. Rev. Neurosci*. **8**, 547–558. (doi:10.1038/nrn2152)
22. D'Ausilio A, Brunetti R, Delogu F, Santonico C, Belardinelli MO. 2010 How and when auditory action effects impair motor performance. *Exp. Brain Res*. **201**, 323–330. (doi:10.1007/s00221-009-2044-6)
23. Haueisen J, Knösche TR. 2001 Involuntary motor activity in pianists evoked by music perception. *J. Cogn. Neurosci*. **13**, 786–792. (doi:10.1162/08989290152541449)
24. Lotze M, Scheler G, Tan H-RM, Braun C, Birbaumer N. 2003 The musician's brain: functional imaging of amateurs and professionals during performance and imagery. *Neuroimage* **20**, 1817–1829. (doi:10.1016/j.neuroimage.2003.07.018)
25. Bangert M, Altenmüller EO. 2003 Mapping perception to action in piano practice: a longitudinal DC-EEG study. *BMC Neurosci*. **4**, 26. (doi:10.1186/1471-2202-4-26)
26. Lahav A, Saltzman E, Schlaug G. 2007 Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *J. Neurosci*. **27**, 308–314. (doi:10.1523/JNEUROSCI.4822-06.2007)
27. D'Ausilio A, Altenmüller E, Olivetti Belardinelli M, Lotze M. 2006 Cross-modal plasticity of the motor cortex while listening to a rehearsed musical piece. *Eur. J. Neurosci*. **24**, 955–958. (doi:10.1111/j.1460-9568.2006.04960.x)
28. Witek MA, Clarke EF, Wallentin M, Kringelbach ML, Vuust P. 2014 Syncopation, body-movement and pleasure in groove music. *PLoS ONE* **9**, e94446. (doi:10.1371/journal.pone.0094446)
29. Stupacher J, Hove MJ, Novembre G, Schütz-Bosbach S, Keller PE. 2013 Musical groove modulates motor cortex excitability: a TMS investigation. *Brain Cogn*. **82**, 127–136. (doi:10.1016/j.bandc.2013.03.003)
30. Candidi M, Sacheli LM, Mega I, Aglioti SM. 2014 Somatotopic mapping of piano fingering errors in sensorimotor experts: TMS studies in pianists and visually trained musically naïves. *Cereb. Cortex* **24**, 435–443. (doi:10.1093/cercor/bhs325)
31. Novembre G, Ticini LF, Schütz-Bosbach S, Keller PE. 2012 Distinguishing self and other in joint action. Evidence from a musical paradigm. *Cereb. Cortex* **22**, 2894–2903. (doi:10.1093/cercor/bhr364)
32. Novembre G, Ticini LF, Schütz-Bosbach S, Keller PE. 2014 Motor simulation and the coordination of self and other in real-time joint action. *Soc. Cogn. Affect. Neurosci*. **9**, 1062–1068. (doi:10.1093/scan/nst086)
33. Hadley LV, Novembre G, Keller PE, Pickering MJ. 2015 Causal role of motor simulation in turn-taking behavior. *J. Neurosci*. **35**, 16 516–16 520. (doi:10.1523/JNEUROSCI.1850-15.2015)
34. D'Ausilio A. 2007 The role of the mirror system in mapping complex sounds into actions. *J. Neurosci*. **27**, 5847–5848. (doi:10.1523/JNEUROSCI.0979-07.2007)
35. D'Ausilio A. 2009 Mirror-like mechanisms and music. *Sci. World J*. **9**, 1415–1422. (doi:10.1100/tsw.2009.160)
36. Grafton ST, Hamilton AFC. 2007 Evidence for a distributed hierarchy of action representation in the brain. *Hum. Mov. Sci*. **26**, 590–616. (doi:10.1016/j.humov.2007.05.009)
37. Wolpert DM, Doya K, Kawato M. 2003 A unifying computational framework for motor control and social interaction. *Phil. Trans. R. Soc. B* **358**, 593–602. (doi:10.1098/rstb.2002.1238)
38. Newman-Norlund RD, van Schie HT, van Zuijlen AM, Bekkering H. 2007 The mirror neuron system is more active during complementary compared with imitative action. *Nat. Neurosci*. **10**, 817–818. (doi:10.1038/nn1911)
39. Brinner B. 1995 *Knowing music, making music*. Chicago, IL: University of Chicago Press.
40. Tal-Shmotkin M, Gilboa A. 2013 Do behaviors of string quartet ensembles represent self-managed teams? *Team Perform. Manage.* **19**, 57–71. (doi:10.1108/13527591311312097)
41. Gilboa A, Tal-Shmotkin M. 2010 String quartets as self-managed teams: an interdisciplinary perspective. *Psychol. Music* **40**, 19–41. (doi:10.1177/0305735610377593)
42. Cohen SG, Ledford GE, Spreitzer GM. 1996 A predictive model of self-managing work team effectiveness. *Hum. Relat*. **49**, 643–676. (doi:10.1177/001872679604900506)
43. Davidson J, Good JMM. 2002 Social and musical co-ordination between members of a string quartet: an exploratory study. *Psychol. Music* **30**, 186–201. (doi:10.1177/0305735602302005)
44. Murnighan JK, Conlon DE. 1991 The dynamics of intense work groups: a study of British string quartets. *Admin. Sci. Quart*. **36**, 165–186. (doi:10.2307/2393352)
45. King EC. 2006 The roles of student musicians in quartet rehearsals. *Psychol. Music* **34**, 262–282. (doi:10.1177/0305735606061855)
46. Seddon F, Biasutti M. 2009 Modes of communication between members of a string quartet. *Small Group Res*. **40**, 115–137. (doi:10.1177/1046496408329277)
47. Schiavio A, Høffding S. 2015 Playing together without communicating? A pre-reflective and enactive account of joint musical performance. *Music. Sci*. 366–388.
48. Butterworth T. 1990 Detroit string quartet. In *Groups that work and those that don't* (ed. JR Hackman), pp. 207–224. San Francisco, CA: Jossey-Bass.
49. Keller P, Appel M. 2010 Individual differences, auditory imagery, and the coordination of body movements and sounds in musical ensembles. *Music Percept*. **28**, 27–46. (doi:10.1525/mp.2010.28.1.27)
50. Goebl W, Palmer C. 2009 Synchronization of timing and motion among performing musicians. *Music Percept*. **26**, 427–438. (doi:10.1525/mp.2009.26.5.427)
51. Berliner P. 1994 *Thinking in jazz: the infinite art of improvisation*. Chicago, IL: University of Chicago Press.
52. Schütz A. 1951 Making music together: a study in social relationship. *Soc. Res*. **18**, 76–97.
53. Fuller J. 1988 How you can hit a hot lick in business. Star Tribune 4–6E.
54. Keller P. 2013 Musical ensemble performance: a theoretical framework and empirical findings on interpersonal coordination. In *Proc. the Int. Symp. on Performance Science 2013* (eds A Williamon, W Goebl), pp. 271–285. Brussels, Belgium: European Association of Conservatoires (AEC).
55. Keller P. 2014 Ensemble performance: interpersonal alignment of musical expression. In *Expressiveness in music performance: Empirical approaches across styles and cultures* (eds D Fabian, R Timmers, E Schubert), pp. 260–282. Oxford, UK: Oxford University Press.
56. Moore GP, Chen J. 2010 Timings and interactions of skilled musicians. *Biol. Cybern*. **103**, 401–414. (doi:10.1007/s00422-010-0407-5)
57. Glowinski D, Mancini M, Cowie R, Camurri A, Chiorri C, Doherty C. 2013 The movements made by performers in a skilled quartet: a distinctive pattern, and the function that it serves. *Front. Psychol*. **4**, 841. (doi:10.3389/fpsyg.2013.00841)
58. Richman J, Moorman J. 2000 Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circ. Physiol.* **278**, H2039.
59. Govindan RB, Wilson JD, Eswaran H, Lowery CL, Preißl H. 2007 Revisiting sample entropy analysis. *Physica A* **376**, 158–164. (doi:10.1016/j.physa.2006.10.077)
60. Wing AM, Endo S, Bradbury A, Vorberg D. 2014 Optimal feedback correction in string quartet synchronization. *J. R. Soc. Interface* **11**, 20131125. (doi:10.1098/rsif.2013.1125)
61. Vorberg D, Wing AM. 1996 Modeling variability and dependence in timing. In *Handbook of perception and action*, vol. 2 (eds H Heuer, S Keele), pp. 181–262. New York, NY: Academic Press.
62. Vorberg D, Schulze H-H. 2002 Linear phase correction in synchronization: predictions, parameter estimation, and simulations. *J. Math. Psychol*. **46**, 56–87. (doi:10.1006/jmps.2001.1375)
63. Badino L, D'Ausilio A, Glowinski D, Camurri A, Fadiga L. 2014 Sensorimotor communication in professional quartets. *Neuropsychologia* **55**, 98–104. (doi:10.1016/j.neuropsychologia.2013.11.012)

64. Granger CWJ. 1969 Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* **37**, 424–438. (doi:10.2307/1912791)
65. Papiotis P, Marchini M, Perez-Carrillo A, Maestre E. 2014 Measuring ensemble interdependence in a string quartet through analysis of multidimensional performance data. *Front. Psychol*. **5**, 963. (doi:10.3389/fpsyg.2014.00963)
66. Moddemeijer R. 1989 On estimation of entropy and mutual information of continuous distributions. *Signal Process*. **16**, 233–246. (doi:10.1016/0165-1684(89)90132-1)
67. Cover T, Thomas J. 2006 *Elements of information theory*, 2nd edn. Hoboken, NJ: Wiley-Interscience.
68. Lehnertz K. 2011 Assessing directed interactions from neurophysiological signals—an overview. *Physiol. Meas*. **32**, 1715–1724. (doi:10.1088/0967-3334/32/11/R01)
69. Chicharro D, Andrzejak R. 2009 Reliable detection of directional couplings using rank statistics. *Phys. Rev. E* **80**, 026217. (doi:10.1103/PhysRevE.80.026217)
70. Varni G, Volpe G, Camurri A. 2010 A system for real-time multimodal analysis of nonverbal affective social interaction in user-centric media. *IEEE Trans. Multimedia* **12**, 576–590. (doi:10.1109/TMM.2010.2052592)
71. Marwan N, Romano MC, Thiel M, Kurths J. 2007 Recurrence plots for the analysis of complex systems. *Phys. Rep*. **438**, 237–329. (doi:10.1016/j.physrep.2006.11.001)
72. Keller P. 2008 Joint action in music performance. In *Enacting intersubjectivity: a cognitive and social perspective to the study of interactions* (eds F Morganti, A Carassa, G Riva), pp. 205–221. Amsterdam, The Netherlands: IOS Press.
73. Orio N. 2006 Music retrieval: a tutorial and review. *Found. Trends Inf. Retrieval* **1**, 1–90. (doi:10.1561/1500000002)
74. Abesser J, Lartillot O, Dittmar C, Eerola T, Schuller G. 2011 Modeling musical attributes to characterize ensemble recordings using rhythmic audio features. In *Proceedings of the IEEE International Conference*, May 22–27, 2011, Prague Congress Centre, Prague, Czech Republic, pp. 189–192. Fraunhofer IDMT, Ilmenau, Germany. (doi:10.1109/ICASSP.2011.5946372)
75. Devaney J, Mandel MI, Fujinaga I. 2012 A study of intonation in three-part singing using the automatic music performance analysis and comparison toolkit (AMPACT). In *Proc. 13th Int. Soc. Music Information Retrieval Conference* (eds F Gouyon, P Herrera, LG Martins, M Mülle), pp. 511–516. Porto, Portugal: FEUP Edições.
76. Makris I, Mullet E. 2009 A systematic inventory of motives for becoming an orchestra conductor: a preliminary study. *Psychol. Music* **37**, 443–458. (doi:10.1177/0305735608100373)
77. Matthews WK, Kitsantas A. 2013 The role of the conductor's goal orientation and use of shared performance cues on collegiate instrumentalists' motivational beliefs and performance in large musical ensembles. *Psychol. Music* **41**, 630–646. (doi:10.1177/0305735612441738)
78. Luck G, Toiviainen P, Thompson MR. 2010 Perception of expression in conductors' gestures: a continuous response study. *Music Percept*. **28**, 47–57. (doi:10.1525/mp.2010.28.1.47)
79. Fredrickson WE. 1994 Band musicians' performance and eye contact as influenced by loss of a visual and/or aural stimulus. *J. Res. Music Ed*. **42**, 306–317. (doi:10.2307/3345738)
80. Luck G, Sloboda JA. 2007 An investigation of musicians' synchronization with traditional conducting beat patterns. *Music Perf. Res*. **1**, 26–46.
81. Wöllner C. 2008 Which part of the conductor's body conveys most expressive information? A spatial occlusion approach. *Music. Sci*. **12**, 249–272. (doi:10.1177/102986490801200204)
82. Bigand E, Lalitte P, Lerdahl F, Boucheix J-M, Gérard Y, Pozzo T. 2010 Looking into the eyes of a conductor performing Lerdahl's 'Time after time'. *Music. Sci*. **14**, 275–294. (doi:10.1177/10298649100140S215)
83. 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**, e35757. (doi:10.1371/journal.pone.0035757)
84. Gnecco G, Glowinski D, Camurri A, Sanguineti M. 2014 On the detection of the level of attention in an orchestra through head movements. *Int. J. Arts Technol.* **7**, 316–338. (doi:10.1504/IJART.2014.066452)
85. Clayton MRL. 2007 Observing entrainment in music performance: video-based observational analysis of Indian musicians' *tanpura* playing and beat marking. *Music. Sci*. **11**, 27–59. (doi:10.1177/102986490701100102)
86. Lucas G, Clayton MRL, Leante L. 2011 Inter-group entrainment in Afro-Brazilian Congado ritual. *Empir. Musicol. Rev.* **6**, 75–102.
87. Alborno P, Keller P, Clayton M, Volpe G, Camurri A. 2015 Automated video analysis of interpersonal entrainment in Indian music performance. In *Proceedings of the 7th International ICST Conference on Intelligent Technologies for Interactive Entertainment* (Intetain 2015), 10–12 June, 2015. Torino, Italy.
88. Babiloni C *et al*. 2012 Brains 'in concert': frontal oscillatory alpha rhythms and empathy in professional musicians. *Neuroimage* **60**, 105–116. (doi:10.1016/j.neuroimage.2011.12.008)
89. Babiloni C *et al*. 2011 Simultaneous recording of electroencephalographic data in musicians playing in ensemble. *Cortex* **47**, 1082–1090. (doi:10.1016/j.cortex.2011.05.006)
90. Vinciarelli A, Pantic M, Bourlard H. 2009 Social signal processing: survey of an emerging domain. *Image Vision Comput*. **27**, 1743–1759. (doi:10.1016/j.imavis.2008.11.007)