doi: 10.5821/conference-9788419184849.10

ANALYZING DIRECTIONALITY OF INFLUENCE AMONG ENSEMBLE MUSICIANS USING GRANGER CAUSALITY

Sanket Rajeev SABHARWAL[1a], Arianna MUSSO[a], Matthew BREADEN[b], Eva RICCOMAGNO[a], Antonio CAMURRI[a], Peter E. KELLER[c]

[a] *University of Genova, Italy*

[b] *MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Australia*

[c] *Center for Music in the Brain, Department of Clinical Medicine, Aarhus University & The Royal Academy of Music Aarhus/Aalborg, Denmark*

ABSTRACT

In small musical groups, performers can seem to coordinate their movements almost effortlessly in remarkable exhibits of joint action and entrainment. To achieve a common musical goal, co-performers interact and communicate using non-verbal means such as upper-body movements, and particularly head motion. Studying these phenomena in naturalistic contexts can be challenging since most techniques make use of motion capture technologies that can be intrusive and costly. To investigate an alternative method, we analyze video recordings of a professional instrumental ensemble by extracting trajectory information using pose estimation algorithms. We examine Kansei perspectives such as the analysis of non-verbal expression conveyed by bodily movements and gestures, and test for causal relationships and directed influence between performers using the Granger Causality method. We compute weighted probabilities representing the likelihood that each performer Granger Causes co-performers' movements. Effects of different aspects of musical textures were examined and results indicated stronger directionality for homophonic textures (clear melodic leader) than polyphonic (ambiguous leadership).

Keywords: *Kansei information processing, Social interaction, Entrainment, Granger causality, Joint actions.*

[1] Corresponding author: sabharwalsanket@gmail.com

1 INTRODUCTION

Temporal data now appear everywhere, and this statement is getting increasingly true with the availability of new modalities, recording devices, and innovations. With recent advancements, studying joint human activities is becoming increasingly accessible, and thus helps with investigating perception and actions among individual and small group setups.

Coordinating tasks is an essential aspect of our lives, and we tend to demonstrate remarkable examples of cognitive and movement-related synchrony to achieve shared task objectives. Humans have a presumably innate capability to entrain their movements, and this is particularly seen in artistic activities such as music and dance where we tend to exhibit intrapersonal and interpersonal entrainment (Clayton et al., 2005; Miyata et al., 2017). The evolutionary process has rendered humans to indulge in social interactions where we may use different communicative capabilities such as voice tonal quality, facial expressions, and bodily gestures, along with cognitive capabilities of perceptual sensitivity and heightened attention to achieve common goals in joint actions (Knoblich & Sebanz, 2006; Tomasello & Vaish, 2013).

In social situations, it is but natural for humans to assume different roles (leader or follower) during joint actions (Chauvigné & Brown, 2018). This can be observed in situations where two people coordinate their movement to lift and move a large heavy object while providing each other with verbal cues. These cues can also be non-verbal, as seen in artistic activities such as dance or musical performances where participants entrain their movements, assume the role of a leader or follower depending on the structural task (Keller et al., 2014; Noy et al., 2011), and adapt to each others' actions at different temporal scales (Chang et al., 2017; Kirschner & Tomasello, 2009; Noy et al., 2011; Shockley et al., 2003). Music ensembles thus serve as ideal test-beds to study leader-follower relationships in joint actions. They tend to act as self-regulated groups that mutually contribute to commonly shared goals of delivering a flawless performance by avoiding errors. With the objective to fulfill the aesthetic and technical goals of a musical composition, ensemble groups members exhibit non-verbal cues such as synchronized bodily movements, especially upper body movements among string quartets (e.g., Keller & Rieger, 2009). During the collective process of co-creation of music, these movements eventually lead to synchronized states where a musical group slowly converges into a single organism (Camurri et al., 2016; Varni et al., 2010).

In this paper, we investigate such subtler, "Kansei"-level, qualitative features of joint actions, particularly the directionality of influence of a leader-follower relationship that emerges during a music performance, which is often communicated using ancillary body movements such as head and body sway (Alborno et al., 2019; Bishop et al., 2021; Chang et al., 2019; Hilt et al., 2019). Based on previous studies (Badino et al., 2014; Hilt et al., 2019), Granger causality (GC) is utilized to quantify the directionality of influence by making use of trajectory information of the head motion (Granger, 1969). These data are extracted using pose estimation algorithms on video sequences of a professional musical ensemble's performance. Unlike previous methods that made use of marker-based motion capture data (MoCap) (Badino et al., 2014; Chang et al., 2017; D'Ausilio et al., 2012; Hilt et al., 2019), we used video recordings with marker-less techniques to localize body coordinates of interest. An important contribution of this study is to propose a conceptual approach to examine "Kansei"-level information by investigating directional flow of non-verbal communication using only video sequences, thus allowing the examination of musicians' movements and associated interactions in naturalistic settings. We also use this technique to then address research questions about the effects of musical structure on leadership dynamics.

This paper is organized as follows: in Section 2, we highlight the research objectives and present relevant studies that have examined leadership dynamics in social situations and on the associated effects of musical textures on leader-follower relationships; Section 3 begins by describing our conceptual approach and methodology to measure directionality of influence, and this is followed by presenting the dataset; Section 4 we present statistical results and in Section 5 we discuss results, conclude the paper, and highlight the limitations and possible future activities.

2 RESEARCH OBJECTIVES AND RELATED WORK

In this research, our objectives are twofold:

1. Propose a conceptual approach and framework for extracting "Kansei"-level features for the analysis of leadership dynamics in small group interactions; and
2. Utilize such an approach to investigate the effects of musical texture and directionality of coupling

In this section we summarize the background and related work with regards to the above-mentioned objectives of this study, i.e., (i) existing methods adopted to analyze leadership dynamics in joint musical actions and (ii) studies examining causal relationships of musical textures.

2.1 Social interaction and leadership in joint musical actions

Humans tend to coordinate tasks with each other almost effortlessly, especially in socially engaging situations such as lifting heavy objects together or jointly playing drums where the limbs' movements eventually synchronize to a common beat (Wallin et al., 2001). In the case of musical ensembles, a shared state of synchronization, often termed entrainment, indicates a spatiotemporal coordination between two or more individuals (Phillips-Silver & Keller, 2012), and is observed among musical performers who engage and communicate with each other non-verbally to achieve the technical and artistic objectives of a musical piece (Clark, 1996; Keller et al., 2016; Sebanz et al., 2006). During such joint actions, the importance of leadership in group dynamics is often highlighted for the successful co-creation of sound (Murnighan & Conlon, 1991; Timmers et al., 2014).

Looking at four relevant studies on GC in musical ensembles, one analyzed position (Chang et al., 2017) and three analyzed acceleration time-series data (Bishop et al., 2019; D'Ausilio et al., 2012; Hilt et al., 2019). Using GC, we can infer the causal relationships between two co-performers' movement-related time series, to check if one (leader) stimulates movement in another (follower) (Glowinski et al., 2012). Following Chang et al. (2017), we use position time-series information, and in addition to the anterior-posterior head sway, we also made use of proximal and distal head movements. Previous research in small group interactions have suggested that the coordination of upper-body movements, particularly head movements, are correlated with the onset of sounds, showcasing that both audio and visual cues are utilized to communicate during music performances (Bishop & Goebl, 2018; Ragert et al., 2013) and head movement coordination is closely linked with a higher sense of connectedness (Latif et al., 2014; Marsh et al., 2009). Additionally, it has been found that the anterior-posterior sway of upper-body movements is informative about leader-follower relationships (Keller & Appel, 2010).

While our proposed approach specifically addresses musical ensembles, it is in essence being utilized to study leader-follower relationships in small group interactions. Musical ensembles happen to be valuable experimental setups to study subtler aspects of non-verbal communication in small groups, and an added benefit is that, as researchers, we can set strong experimental controls (D'Ausilio et al., 2015).

2.2 Effects of musical texture

In the present study, we address how the variation in directionality of coupling relates to the musical texture, a structural aspect of music related to the complexity of the relationship between multiple simultaneously sounding parts. While it is not definitively clear how textural changes can affect group dynamics and eventual coordination, based on previous research (Novembre et al., 2015; Noy et al., 2011; Sabharwal et al., 2022; Varlet et al., 2020), we hypothesize that our measurement of the influence in directionality of coupling would be higher in homophonic textures than in polyphonic textures. Homophonic textures are those where there is a clear melodic leader while in polyphonic textures the leadership is distributed among co-performers.

3 CONCEPTUAL FRAMEWORK, DATA AND METHODOLOGY

In a previous study (Sabharwal et al., 2022), we proposed huSync (Human Sync), a multimodal computational model and system to study the small group interactions that entail non-verbal social communicative behavior. We utilized huSync to quantify interpersonal coupling in a small musical group and revealed how dyadic synchronization varied as a function of textural demands and position within the musical phrase. As an extension of our previous work, and to address the research objectives of this study, we make an attempt to quantify mutual interaction, and leadership and study the effects of musical structure on the directionality of coupling by making use of relatively long timescales of upper-body movements such as head sway (Alborno et al., 2019; Chang et al., 2019; Hilt et al., 2019).

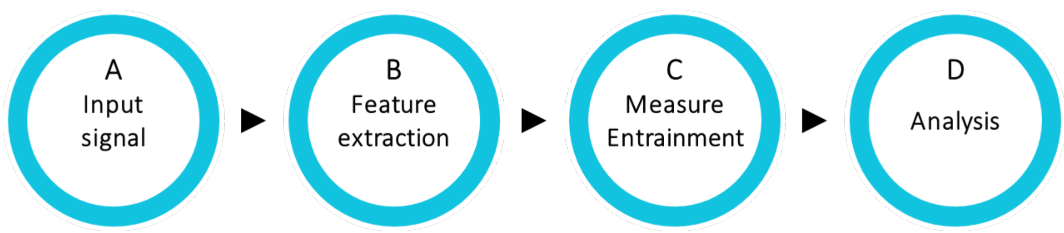

Figure 1 An illustration of our experimental approach

From the computational model shared in Sabharwal et al. (2022) (Ref. Fig. 1), we derive our experimental approach which consists of the following four steps:

1. **Input signal**: In the first step (Fig. 1(A)) we read the video signals from standard video recordings and pre-process using the pose estimation algorithm AlphaPose (Fang et al., 2018), to obtain a json file consisting of trajectory information of (x,y) body part locations sampled at 30Hz.
2. **Feature extraction of relevant key-point(s)**: On observing the outputs available using AlphaPose, the best representation of the head from the available face landmarks was the nose key-point, and here we select it as a cue on head sway. The nose trajectory data are then utilized to quantify directionality of influence among performers in the next block (Fig. 1(B)).
3. **Measurement of entrainment**: In the third block (Fig. 1(C)), we measure the directionality of information flow by applying GC using the Statsmodels Python module (Seabold & Perktold, 2010), on time-series data of the X and Y coordinates separately for all possible dyadic pairs. For the viola player (Fig. 2, center), due to the seating position and being in the center of the camera view, an anterior-posterior sway is not effectively tracked, and thus we also use the Y coordinate to track proximal and distal movements, which were noted to be reliable during our feasibility tests. On performing the GC test to check if time-series A causes B, and if $p<.05$ we reject the null hypothesis and infer that the past values of A have a statistically significant effect on the current value of B – thereby demonstrating a causal relationship between the two time-series. In musical performances, as in most behaviors, it is common to have a delay in between a stimulus and response (Chang et al., 2017, p. 20; Meals, 2020), and to address this issue, we performed GC tests for a lag of ~1 sec and used multiple lag lengths up to 30 (equal to the sampling rate), for both the X and Y coordinates of the nose key-point separately for each pair.
4. **Analysis**: In the fourth step (Fig. 1(D)) GC results obtained are mapped in a tabular layout for further analysis. Specifically, for each phrase analyzed, a table is made with the column values indicating the Granger causing time-series, with 1 representing the left-most musician (violin) and 5 representing the right-most (clarinet). We assign a '1' for cells where we detect GC (irrespective of lag) and assign '0' for other cells where the time-series were either not stationary or where $p>.05$. Fig. 3 represents the mapping process adopted to assign binary values to the GC results. Here, 'NC' implies 'Not calculated' since either one of the time-series was not stationary, and was also assigned '0', while NA implies 'Not applicable'.

	X Coordinate				
	Granger Causing				
	1	2	3	4	5
1	NA	Yes	NC	Yes	Yes
2	Yes	NA	NC	Yes	Yes
3	NC	NC	NA	NC	NC
4	Yes	Yes	NC	NA	Yes
5	Yes	Yes	NC	Yes	NA

→

	X Coordinate				
	Granger Causing				
	1	2	3	4	5
1	NA	1	0	1	1
2	1	NA	0	1	1
3	0	0	NA	0	0
4	1	1	0	NA	1
5	1	1	0	1	NA

(A)

	Y Coordinate				
	Granger Causing				
	1	2	3	4	5
1	NA	Yes	NC	Yes	Yes
2	Yes	NA	NC	Yes	Yes
3	NC	NC	NA	NC	NC
4	Yes	Yes	NC	NA	Yes
5	Yes	Yes	NC	Yes	NA

→

	Y Coordinate				
	Granger Causing				
	1	2	3	4	5
1	NA	1	0	1	1
2	1	NA	0	1	1
3	0	0	NA	0	0
4	1	1	0	NA	1
5	1	1	0	1	NA

(B)

Figure 2 Mapping process of GC results to binary values

This framework was used on a dataset consisting of performance recordings by the Omega Ensemble, a professional chamber music group from Australia. For our study, we performed an analysis on recordings (Fig. 2) of the Clarinet Quintet in B minor (Op. 115) written in 1891 by Johannes Brahms (1833-1897).

Figure 3 An image from the performance of the Brahms Clarinet Quintet (Left) and output available with tracked key-points using a pose estimation algorithm (Right)

Videos were recorded using a front camera, and split into parts based on textural annotations (homophonic and polyphonic) that were made using ELAN (ELAN, 2022) based on a musicological analysis of the published score. In Table 1 we report the number of phrases that were selected for our study. We made use of an equal number of homophonic and polyphonic phrases to have balanced textural classifications, and their minimum, maximum, median, and average duration.

Table 1 Summary of the complete dataset and selected phrases for our experiments.

	Complete Dataset					Selected Phrases				
	Duration (s)				Count	Duration (s)				Count
Texture	Minimum	Maximum	Median	Average		Minimum	Maximum	Median	Average	
Homophonic	15.03	38.20	19.74	21.57	27	16.75	34.87	21.60	23.51	12
Polyphonic	15.49	33.08	23.10	23.53	20	15.49	27.55	20.16	21.11	12

4 RESULTS

A two-way analysis of variance (ANOVA) was conducted on the GC values yielded by the analysis procedure (Fig. 1(D)), and statistical tests were run in jamovi version 1.6.23 (The jamovi project, 2021). In line with our hypothesis, the reported estimated marginal means in Fig. 4 indicates that the mean value of homophonic textures, where we observe distinct melodic leadership, is higher as compared to the polyphonic texture where leadership is rather ambiguous.

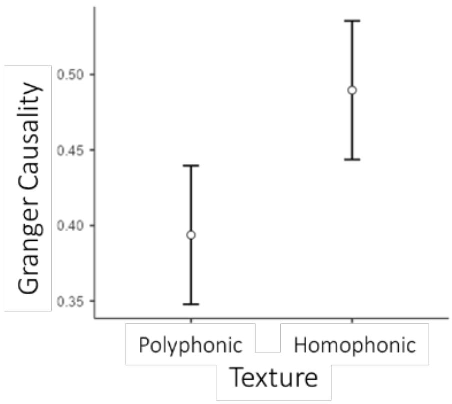

Figure 4 A plot showing the mean values of GC for homophonic and polyphonic textures

An approach for ANOVA using binary values was adopted (Luepsen, 2021), and we entered in it GC results for all combinations of dyadic pairs. Head Sway (Nose trajectory data as X and Y coordinates) was included as a within subjects factor, and Texture (Homophonic, Polyphonic) and Pair (each pair represented as M1-M2, where M1 granger causes M2) as between subjects factors. The ANOVA results are illustrated in Table 2, where values revealing statistical significance have been highlighted in green.

Table 2 ANOVA results for Between and Within Subjects Effects for the Brahms concert

Between Subjects Effects					
	Sum of Squares	df	Mean Square	F	p
Pair	10.15	19	0.534	2.041	0.006
Texture	2.2	1	2.204	8.421	0.004
Pair ✻ Texture	3.21	19	0.169	0.646	0.871
Residual	115.17	440	0.262		

Within Subjects Effects					
	Sum of Squares	df	Mean Square	F	p
Head Sway	4	1	4.004	19.047	<.001
Head Sway ✻ Pair	4.16	19	0.219	1.042	0.410
Head Sway ✻ Texture	1.67	1	1.667	7.928	0.005
Head Sway ✻ Pair ✻ Texture	3.67	19	0.193	0.918	0.561
Residual	92.5	440	0.21		

Results revealed statistically significant main effects of Texture, $F(1,440) = 8.421$, $p = 0.004$, Pair, $F(19,440) = 2.041$, $p = 0.006$, and Head Sway, $F(1,440) = 19.047$, $p<.001$. Additionally, a significant two-way interaction between Head Sway and Texture was also observed, $F(1,440) = 7.928$, $p=0.005$. Head Sway (X and Y coordinates) was included in the ANOVA to account for variance related to movement direction, but this variable was not of theoretical interest and the results will not be presented for the separate dimensions. While these are early findings, results for the texture effect confirms stronger evidence for directional coupling in homophonic textures.

5 DISCUSSION AND CONCLUSION

The contributions of this work are two-fold. On one hand, we present our experimental approach to study the directionality of coupling among small group setups, and we do so using non-intrusive methods. On the other hand, we test our approach on ensemble music performances in an attempt to extract finer subtleties, and aspects, of non-verbal expression. We analyze how head sway influences unfolding leadership dynamics among co-performers, and that stronger directional influence is observed in homophonic textures characterized by the presence of a melodic leader) as compared to a more egalitarian polyphonic textures. A key benefit of this approach is the use of video sequences as input signals which enables experimentation with naturalistic movement patterns.

We would like to highlight some limiting factors observed in our approach. While alternative tracking methods such as marker-based systems are expensive and can cause discomfort among participants, the data available are relatively less noisy as compared to marker-less techniques. There are technical challenges to overcome with marker-less methods such as dependency on frame resolution, occlusion, and variations in illumination, to name a few. Another limitation observed is that the data need to be stationary when using GC. If we apply any techniques to resolve this, we modify the underlying nature of the time-series from non-stationary to stationary, and this can affect the outcome of results. Previous studies (Pearl, 2000) suggest that GC could be used for statistical studies but making use of it for causal relationships is questionable. Additionally, it is still unclear how textural variations affect group entrainment and dynamics, and the preliminary findings of this study are not appropriate for making conclusive or general statements.

In the future, we aim to further investigate the effects of musical structure and phrase position on leadership dynamics in small group setups, and this includes experimenting with more data and validating the results with analysis of acoustic features. As highlighted in Sabharwal et al. (2022) as well, this approach and framework provides a dependable and an unintrusive method in comparison to current methodologies for studying expressive, or “Kansei”-like, aspects of human body movements and their associated qualities.

ACKNOWLEDGMENTS

The authors would like to thank the members of the Omega Ensemble, especially Artistic Director and Founder David Rowden. The authors also thank videographer Bruce Terry for making available the unprocessed video files and Stephanie Laun for assistance with video annotation. This research reported in the paper was part of the H2020 FETPROACTIVE EnTimeMent Project, supported by funding from the European Union’s Horizon 2020 research and innovation program under grant agreement No 824160.

REFERENCES

Alborno, P., Volpe, G., Mancini, M., Niewiadomski, R., Piana, S., & Camurri, A. (2019). The Multi-Event-Class Synchronization (MECS) Algorithm. *ArXiv:1903.09530 [Cs]*. http://arxiv.org/abs/1903.09530

Badino, L., D'Ausilio, A., Glowinski, D., Camurri, A., & Fadiga, L. (2014). Sensorimotor communication in professional quartets. *Neuropsychologia*, *55*, 98–104. https://doi.org/10.1016/j.neuropsychologia.2013.11.012

Bishop, L., Cancino-Chacón, C., & Goebl, W. (2019). Moving to Communicate, Moving to InteractPatterns of Body Motion in Musical Duo Performance. *Music Perception*, *37*(1), 1–25. https://doi.org/10.1525/mp.2019.37.1.1

Bishop, L., & Goebl, W. (2018). Communication for coordination: Gesture kinematics and conventionality affect synchronization success in piano duos. *Psychological Research*, *82*(6), 1177–1194. https://doi.org/10.1007/s00426-017-0893-3

Bishop, L., González Sánchez, V., Laeng, B., Jensenius, A. R., & Høffding, S. (2021). Move like everyone is watching: Social context affects head motion and gaze in string quartet performance. *Journal of New Music Research*, *50*(4), 392–412. https://doi.org/10.1080/09298215.2021.1977338

Camurri, A., Volpe, G., Piana, S., Mancini, M., Niewiadomski, R., Ferrari, N., & Canepa, C. (2016). *The Dancer in the Eye: Towards a Multi-Layered Computational Framework of Qualities in Movement*. 1–7. https://doi.org/10.1145/2948910.2948927

Chang, A., Kragness, H. E., Livingstone, S. R., Bosnyak, D. J., & Trainor, L. J. (2019). Body sway reflects joint emotional expression in music ensemble performance. *Scientific Reports*, *9*(1), 205. https://doi.org/10.1038/s41598-018-36358-4

Chang, A., Livingstone, S. R., Bosnyak, D. J., & Trainor, L. J. (2017). Body sway reflects leadership in joint music performance. *Proceedings of the National Academy of Sciences*, *114*(21), E4134–E4141. https://doi.org/10.1073/pnas.1617657114

Chauvigné, L. A. S., & Brown, S. (2018). Role-Specific Brain Activations in Leaders and Followers During Joint Action. *Frontiers in Human Neuroscience*, *12*, 401. https://doi.org/10.3389/fnhum.2018.00401

Clark, H. H. (1996). *Using Language*. Cambridge University Press.

Clayton, M., Sager, R., & Will, U. (2005). In time with the music: The concept of entrainment and its significance for ethnomusicology. *European Meetings in Ethnomusicology*, *11,* 3–142.

D'Ausilio, A., Badino, L., Li, Y., Tokay, S., Craighero, L., Canto, R., Aloimonos, Y., & Fadiga, L. (2012). Leadership in Orchestra Emerges from the Causal Relationships of Movement Kinematics. *PLOS ONE*, *7*(5), e35757. https://doi.org/10.1371/journal.pone.0035757

D'Ausilio, A., Novembre, G., Fadiga, L., & Keller, P. E. (2015). What can music tell us about social interaction? *Trends in Cognitive Sciences*, *19*(3), 111–114. https://doi.org/10.1016/j.tics.2015.01.005

ELAN, 2022. (2022). *ELAN (Version 6.3) [Computer software]. (2022). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive.* Retrieved from https://archive.mpi.nl/tla/elan

Fang, H.-S., Xie, S., Tai, Y.-W., & Lu, C. (2018). RMPE: Regional Multi-person Pose Estimation. *ArXiv:1612.00137 [Cs]*. http://arxiv.org/abs/1612.00137

Glowinski, D., Badino, L., & D'Ausilio, A. (2012). *Analysis of Leadership in a String Quartet*. 6.

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, *37*(3), 424–438. https://doi.org/10.2307/1912791

Hilt, P. M., Badino, L., D'Ausilio, A., Volpe, G., Tokay, S., Fadiga, L., & Camurri, A. (2019). Multi-layer adaptation of group coordination in musical ensembles. *Scientific Reports*, *9*(1), 5854. https://doi.org/10.1038/s41598-019-42395-4

Keller, P. E., & Appel, M. (2010). Individual Differences, Auditory Imagery, and the Coordination of Body Movements and Sounds in Musical Ensembles. *Music Perception*, *28*(1), 27–46. https://doi.org/10.1525/mp.2010.28.1.27

Keller, P. E., Novembre, G., & Hove, M. J. (2014). Rhythm in joint action: Psychological and neurophysiological mechanisms for real-time interpersonal coordination. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1658). https://doi.org/10.1098/rstb.2013.0394

Keller, P. E., Novembre, G., & Loehr, J. (2016). Musical Ensemble Performance: Representing Self, Other and Joint Action Outcomes. In S. S. Obhi & E. S. Cross (Eds.), *Shared Representations* (pp. 280–310). Cambridge University Press. https://doi.org/10.1017/CBO9781107279353.015

Keller, P. E., & Rieger, M. (2009). Special Issue—Musical Movement and Synchronization. *Music Perception: An Interdisciplinary Journal*, *26*(5), 397–400. https://doi.org/10.1525/mp.2009.26.5.397

Kirschner, S., & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. *Journal of Experimental Child Psychology*, *102*(3), 299–314. https://doi.org/10.1016/j.jecp.2008.07.005

Knoblich, G., & Sebanz, N. (2006). The Social Nature of Perception and Action. *Current Directions in Psychological Science*, *15*(3), 99–104. https://doi.org/10.1111/j.0963-7214.2006.00415.x

Latif, N., Barbosa, A. V., Vatiokiotis-Bateson, E., Castelhano, M. S., & Munhall, K. G. (2014). Movement Coordination during Conversation. *PLoS ONE*, *9*(8), e105036. https://doi.org/10.1371/journal.pone.0105036

Luepsen, H. (2021). ANOVA with binary variables: The F-test and some alternatives. *Communications in Statistics - Simulation and Computation*, *0*(0), 1–25. https://doi.org/10.1080/03610918.2020.1869983

Marsh, K. L., Richardson, M. J., & Schmidt, R. C. (2009). Social connection through joint action and interpersonal coordination. *Topics in Cognitive Science*, *1*(2), 320–339. https://doi.org/10.1111/j.1756-8765.2009.01022.x

Meals, C. D. (2020). The Question of Lag: An Exploration of the Relationship Between Conductor Gesture and Sonic Response in Instrumental Ensembles. *Frontiers in Psychology*, *11*. https://www.frontiersin.org/article/10.3389/fpsyg.2020.573030

Miyata, K., Varlet, M., Miura, A., Kudo, K., & Keller, P. E. (2017). Modulation of individual auditory-motor coordination dynamics through interpersonal visual coupling. *Scientific Reports*, *7*(1), 16220. https://doi.org/10.1038/s41598-017-16151-5

Murnighan, J. K., & Conlon, D. E. (1991). The Dynamics of Intense Work Groups: A Study of British String Quartets. *Administrative Science Quarterly*, *36*(2), 165–186. https://doi.org/10.2307/2393352

Novembre, G., Varlet, M., Muawiyath, S., Stevens, C. J., & Keller, P. E. (2015). The E-music box: An empirical method for exploring the universal capacity for musical production and for social interaction through music. *Royal Society Open Science*, *2*(11), 150286. https://doi.org/10.1098/rsos.150286

Noy, L., Dekel, E., & Alon, U. (2011). The mirror game as a paradigm for studying the dynamics of two people improvising motion together. *Proceedings of the National Academy of Sciences*, *108*(52), 20947–20952. https://doi.org/10.1073/pnas.1108155108

Pearl, J. (2000). *Causality: Models, reasoning and inference* (Vol. 19). Pearl, Judea

Phillips-Silver, J., & Keller, P. (2012). Searching for Roots of Entrainment and Joint Action in Early Musical Interactions. *Frontiers in Human Neuroscience*, *6*. https://www.frontiersin.org/article/10.3389/fnhum.2012.00026

Ragert, M., Schroeder, T., & Keller, P. (2013). Knowing too little or too much: The effects of familiarity with a co-performer's part on interpersonal coordination in musical ensembles. *Frontiers in Psychology*, *4*. https://www.frontiersin.org/article/10.3389/fpsyg.2013.00368

Sabharwal, S. R., Varlet, M., Breaden, M., Volpe, G., Camurri, A., & Keller, P. E. (2022). *HuSync—A model and system for the measure of synchronization in small groups: A case study on musical joint action. *Manuscript submitted**.

Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and Statistical Modeling with Python. *Proceedings of the 9th Python in Science Conference*, 92–96. https://doi.org/10.25080/Majora-92bf1922-011

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, *10*(2), 70–76. https://doi.org/10.1016/j.tics.2005.12.009

Shockley, K., Santana, M.-V., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 326–332. https://doi.org/10.1037/0096-1523.29.2.326

The jamovi project, 2021. (2021). *. Jamovi (Version 1.6) [Computer Software]. Retrieved from https://www.jamovi.org*.

Timmers, R., Endo, S., Bradbury, A., & Wing, A. M. (2014). Synchronization and leadership in string quartet performance: A case study of auditory and visual cues. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00645

Tomasello, M., & Vaish, A. (2013). Origins of Human Cooperation and Morality. *Annual Review of Psychology*, *64*(1), 231–255. https://doi.org/10.1146/annurev-psych-113011-143812

Varlet, M., Nozaradan, S., Nijhuis, P., & Keller, P. E. (2020). Neural tracking and integration of 'self' and 'other' in improvised interpersonal coordination. *NeuroImage*, *206*, 116303. https://doi.org/10.1016/j.neuroimage.2019.116303

Varni, G., Volpe, G., & Camurri, A. (2010). A System for Real-Time Multimodal Analysis of Nonverbal Affective Social Interaction in User-Centric Media. *IEEE Transactions on Multimedia*, *12*(6), 576–590. https://doi.org/10.1109/TMM.2010.2052592

Wallin, N. L., Merker, B., & Brown, S. (2001). *The Origins of Music*. MIT Press.