Title: Measuring Acoustic Social Worlds: Reflections on a Study of Multi-Agent Human Interaction

Word Count: 3842 words

Abstract: 251 words

Authors: Shannon Proksch[1], Majerle Reeves[2], Michael Spivey[1], Ramesh Balasubramaniam[1]

Affiliation: [1]Cognitive & Information Sciences, University of California, Merced; [2]Applied Mathematics, University of California, Merced

Keywords: Social Interaction; Emergence; Dynamical Systems; Music; Interpersonal Coordination

Short Author Bios:

Shannon Proksch is a Ph.D. Candidate in Cognitive and Information Sciences at the University of California, Merced. She received her B.A. in Music from Texas A&M University, Corpus Christi in 2015, and her M.Sc. in Mind, Language and Embodied Cognition from the University of Edinburgh in 2017. Her research blends empirical and philosophical methods from the dynamical systems and predictive processing frameworks to examine the neural, behavioral, and social dynamics of music cognition—from lower-level beat processing to higher-level coordination and social interaction. Her work is supported, in part, by a National Science Foundation Traineeship in Intelligent and Adaptive Systems.

Majerle Reeves is a Ph.D. candidate in the Applied Mathematics department at the University of California, Merced. She has a B.A. in Mathematics and a B.S. in Mechanical Engineering from California State University, Fresno in 2017. Currently, she works on interdisciplinary projects applying her mathematical knowledge to problems in Public Health and

Cognitive Science. Her research focuses on building interpretable machine learning and statistical models to predict suicide death from administrative patient records. Her work is supported, in part, by a National Science Foundation Traineeship in Intelligent and Adaptive Systems.

Michael J. Spivey has a Ph.D. in Brain and Cognitive Sciences from University of Rochester and is Professor of Cognitive Science at the University of California, Merced. After 12 years as a psychology professor at Cornell University, Spivey moved to UC Merced in 2008 to help build their Cognitive & Information Sciences Ph.D. Program. His research uses eye-tracking, computer-mouse tracking, and dynamical systems theory to explore how brain, body, and environment work together to make a mind what it is. In 2010, Spivey received the William Procter Prize for Scientific Achievement from the Sigma Xi Scientific Research Honor Society.

Ramesh Balasubramaniam is a Professor of Cognitive and Information Sciences at the University of California, Merced. He received his PhD from the University of Connecticut in 2000. The primary aim of his research is to understand the organization of human action, with the eventual goal of developing a comprehensive theory of embodied cognition. Dr. Balasubramaniam's research program uses methods from complex dynamical systems, robotics, neuroimaging, and statistical physics to study human movement production at the Cognitive, Neurophysiological and Ecological levels of analysis. Dr. Balasubramaniam’s work is supported by several grants from the National Science Foundation and other funding agencies.

Abstract

When individuals are sharing similar behavioral, physiological, or neural states—that is, when individuals' actions, body states, or brain activity are changing together in time—then a collective interpersonal synergy forms (Riley et al., 2011). Each individual starts to behave together as a member of one large group. Measurement of an interpersonal synergy can be indicative of shared social cognition—of joint participation in co-regulating multiple patterns of activity between two or more agents engaged in a social interaction (De Jaegher, Di Paolo, Gallagher, 2010). In order to examine the dynamics of multiagent groups of people, and the emergence of these interpersonal synergies, scientists generally measure signals emitted by each *individual* to detect correlations between these signals.

However, humans are highly capable of tracking complex behavioral dynamics of multiagent groups in our everyday interactions with the world even without access to individual behavioral, physiological, or neural signals from each individual agent. Our everyday human interactions provide us with access to a shared and co-created acoustic social world. Interactions among members of musical ensembles in particular provide useful insight into the co-creation of acoustic social worlds and the emergence of collective synergies. In the context of 4E cognition and the dynamical systems framework, this chapter provides an examination of the measurement of collective synergies from acoustic signals not of individual musicians, but of the acoustic signal co-created by a musical ensemble as a whole, in performance of a musical work that was composed to highlight the emergence of such a collective synergy.

Introduction

Humans are highly capable of tracking complex behavioral dynamics of crowds in our everyday interactions with the world. Imagine the sounds of a crowded coffee shop. Consider how individuals in that coffee shop might be interacting with each other. There may be some small groups or pairs, and many individuals engaging in small, temporary interactions—but these individuals are not *coordinating* with every other individual in a cohesive “coffee shop group”. They are just a jumble of individuals cohabiting a shared space.

Now, imagine the audience on the floor of a rock concert, cheering or singing along with the artists on stage. Alternatively, imagine the fan section at your favorite sporting event emerging into synchronous chant, or a chorus of resounding “boos”. Imagine how individuals in those large crowds might be interacting with each other. As they cheer, or sing, or chant, or boo, they are sharing similar behavioral states—engaging in similar actions. They are likely sharing similar physiological and neural states as well. These crowds are changing together in time. They are behaving—-coordinating—like one large interdependent group. They have formed a synergy.

The differences between this coffee shop crowd and the rock concert audience or the sporting event crowd are trivially easy for you or I to identify. An uncoordinated group of independent individuals, happening to coexist in a shared space versus a coordinated interdependent group of members of a crowd. We can simply *hear* that these two groups of people sound different. Similarly, we can simply *hear* when we are engaged and participating in a collective synergy while we are interacting in a large group of people. Or perhaps, in a very large musical ensemble.

As scientists in the lab, we are able to measure one or many signals from *every individual* in an interacting or non-interacting group. We can measure those signals and, usually, identify which members of the group are sharing the same *behavioral* state by identifying when their movements synchronize with, or complement, other members of the group. We can measure the pattern of each individual's heartbeats or electrodermal activity (subtle electrical signals conducted by the skin) to identify when individuals are sharing the same *physiological* state. We can even measure precise patterns of brain activity (EEG, or fMRI) to identify when individuals are sharing the same *neural* state (Balconi & Fronda, 2020; Misaki et al., 2021; Schirmer et al., 2021). When individuals are sharing similar behavioral, physiological, or neural states—that is, when individuals' actions, body states, or brain activity are changing together in time—then a collective interpersonal synergy forms (Riley et al., 2011). Each individual starts to behave together as a member of one large group. Measurement of an interpersonal synergy can be indicative of shared social cognition: of joint participation in co-regulating multiple patterns of activity between two or more agents engaged in a social interaction (De Jaegher, Di Paolo, Gallagher, 2010).

However, as humans going about our everyday lives, we don't carefully measure the individual components of our successful social interactions or the emergence of interpersonal synergies in our joint actions to determine whether we are engaged in a participatory instance of social cognition. We don't carefully monitor our own movement and brain patterns or carefully compare each of these many individual signals that we generate to the many individual signals that our friends, coworkers, or other individuals generate as we are interacting in real time. In fact, we don't even have access to individualised movement, body, and brain data of ourselves and the people with whom we interact on a daily basis.

Scale this up to an interaction between four or five people, or even further to a very large audience attending a concert—and this feat is unmanageable for any single human's sensory system. Often, you might lack even visual access to every member of the group you are interacting in. What you do have access to is a shared and co-created acoustic world (albeit from your own unique point of reference). With this in mind, we turn first to the science of interpersonal coordination, and then to an empirical study evaluating the formation of interpersonal synergy within a set of musicians, who join together in co-creating an acoustic musical world.

Interpersonal Coordination

The science of interpersonal coordination has made advances in describing how individuals interact as part of a dyad or a large group by evaluating a multitude of movement, body, and brain signals from each individual in comparison with each other individual.

But what if scientists don’t have access to the vast array of recording devices they rely upon in their lab? What if a scientist wants to study how real groups interact in the wild? Can science identify these same differences in an uncoordinated group of independent individuals versus a coordinated, interacting and interdependent crowd? This is the question we asked in a recent study of multi-agent interaction within a musical ensemble. These musicians performed a piece that was specifically composed so that the musicians first create uncoordinated noise for a period of time on each of their instruments, before joining together into a coordinated joint musical performance. Proksch et al. (2021) wanted to understand how the musicians were changing their acoustic behavior in time, either independently or *inter*dependently, in the two different musical interactions dictated by their musical score. With the understanding that individuals often do not have unobstructed visual access to every member of an interacting group (let alone physiological or neural data), we restricted our dynamical systems analysis to a single measurement of the shared and co-created behavior generated by the musicians—a raw audio file of the ensemble’s acoustic behavior.

The dynamical systems framework in cognitive science allows for the study of the formation of interpersonal synergies. An interpersonal synergy occurs when the movement dynamics of one individual become causally coupled to the movement dynamics of another individual (Riley et al., 2011). This means that the actions of cognitive agents constrain each other, interacting as a single coupled unit. Interpersonal synergies can arise from simple interactions, such as walking through the park engaged in a conversation and finding one has begun walking in step with your conversation partner (Atherton, Sebanz, & Cross, 2019). Subtler interpersonal synergies can arise in conversation when standing still. Even if one cannot see their conversation partner, the mere act of interacting through conversation serves to constrain subtle sway patterns of body movement, such that body movement is distinctly coupled to the movement of the unseen partner (Shockley, Santana, Fowler, 2003).

Perhaps more immediately observable, however, are the interpersonal synergies which we see and hear in musical interactions. Where conversation partners might incidentally fall into step or sway together in time, a pair or a group of musicians co-creating a musical performance are intentionally coordinating their acoustic behavior. It’s important to note here, that coordinating acoustic behavior in order to engage in a successful musical interaction often involves musicians moving their bodies differently from their musical partners. A trombone player will make different movements than the string bass player, and a trumpet player or pianist will make different movements and perhaps even play more notes in the same amount of time compared to the trombonist and bassist. But together, these differing movement dynamics from each musician join to co-create the same shared acoustic social world. The acoustic output of each musician constrains the acoustic output, and motor behavior, of each other musician in the ensemble. In fact, if the low voices (the trombone and the bassist) were to play a continuous drone, one single chord for a prolonged time, then the duet that the trumpet player and pianist improvise together may result in a different ‘performance narrative’ than if the low voices provided a rhythmic bass line. In a study involving duets performed by pairs of skilled pianists, improvising over the unstructured ‘drone’ backing track resulted in increased movement coordination between the two pianists compared to improvisation over the rhythmic bass line (Walton et al., 2018). Specifically, pianists repeated their improvisation partner’s note combinations and head movements in longer sequences when improvising over the drone backing track. Further, listeners rated this performance as more ‘harmonious’ than the improvisation over the structured, rhythmic bass line, with listeners giving higher ‘harmonious’ ratings when the musicians repeated each other’s note combinations for longer sequences of time (Walton et al., 2018).

Experimental setups are typically designed to identify interpersonal synergies by correlating one or more of the movement/body/neural signals from each member of the interacting or non-interacting group. But in principle—once an interpersonal synergy is formed—it should be possible to analyze group behavioral dynamics from *one* single signal measured from that system. This is due to two factors—dimensional compression, and reciprocal compensation (Riley et al., 2011). Dimensional compression within a synergy occurs when the movement of many potentially independent elements (such as the movement of two independent pairs of legs on two independent walking individuals) become coupled so that they move in time together (the two pairs of legs begin to walk in step, as one interdependent walking dyad). Reciprocal compensation, also termed mutual adaptation, describes the ability of movement in one element of a synergy to react to, or adapt to, the movement of another element of the synergy (one member of the walking dyad can adjust their walking speed to ensure they are in step with their walking partner; Riley et al., 2011). The behavior of the musicians in the improvising piano duets we visited earlier exhibited these features of dimensional compression and reciprocal compensation. The movement of two independent pairs of hands, and two independent heads, on two individual musicians became coupled so that they created music in time together, and the musical behavior of each musician reacted to, or adapted to, the musical behavior of the other musician. The listeners, who rated this musical performance, were able to extract an acoustic signal from that system and attune to differences in how the two duetting pianists coordinated their sound and movement (Walton et al., 2018). If these listeners were able to attune to these differences in coordination in two forms of coordinated music making (improvised duets over two different backing tracks) based on a single acoustic signal—the raw audio of the music performance itself—then perhaps this same feat can be scaled up to a multi-agent interaction of a much larger musical ensemble. And if the motor and acoustic behavior of individuals within a much larger musical ensemble are functioning together in time, so as to have the features of dimensional compression and reciprocal compensations necessary to form an interpersonal synergy, then it should be possible to detect that synergy from something as sparse as a raw audio file.

An Empirical Study of Multi-Agent Musical Interaction

We investigated the coordination dynamics of a performance of “Welcome to the Imagination World”, composed by Daisuke Shimizu (Shimizu, 2016) and performed by the

Inagakuen Wind Orchestra (Proksch et al. (2021))[1]. Specifically, we evaluated the acoustic behavior of this musical ensemble using methods from dynamical systems theory of phase space reconstruction and recurrence quantification analysis. These time series analysis methods allow researchers to detect two features of interpersonal synergies discussed above—dimensional compression and reciprocal compensation—and to measure patterns of this synergistic behavior over time. In this case, we were somewhat playing the role of the listeners of the improvising pianists. The difference was, instead of asking whether one could *hear* differences in coordination, the question was whether one might *empirically measure* differences in coordination using those dynamical systems tools. And importantly, can these differences in coordination be measured using only the raw audio signal of the musical performance, without access to individual recordings of each musicians' acoustic output.

The musical performance was divided into two main coordination categories: uncoordinated and coordinated. This uncoordination was in fact a specific feature of the musical composition itself. Shimizu composed this piece to reflect the "arrival and development of a simple fanfare motif into an accomplished work" (windrep.org), beginning with "random ad lib music...free of tempo and as expressive as possible" until the musicians invite the conductor on stage as the horn, tenor, and brass instruments unify into a majestic introduction' (windrep.org). These descriptions are from program notes describing "Welcome to the Imagination World". Listening to a performance of this work, one can easily hear the difference between the uncoordinated improvisations of individual musicians on stage and the coordinated, collective interaction of the musicians as they co-create 'an accomplished work'. Importantly, however, we were also able to measure those differences in coordination dynamics from the raw audio signal.

---

[1] https://www.youtube.com/watch?v=-wJ9ZsgO3QI

The results from the time series analysis make clear that there is a detectable difference between the uncoordinated and coordinated portions of the performance. They tell us that when the musicians began to coordinate their actions, such that the actions of each musician became interdependent on the action of each of the other musicians in the ensemble, they formed a single complex system—a collective interpersonal synergy.

Discussion

What are the implications of measuring these differences in coordination dynamics between the acoustic behavior of non-interacting uncoordinated musicians gathered on a stage versus the acoustic behavior of the same musicians when interacting and coordinated as a cohesive musical ensemble?

The current pragmatic turn in cognitive science toward action-oriented views of cognition (c.f. Engel, Friston & Kragic, 2016) provides a useful explanatory viewpoint for discussing coordination as it relates to multi-agent musical interaction. Specifically, we can interpret the collective coordination which emerged in this musical ensemble in terms of sensorimotor contingency theory under the cognitive framework of enactivism. We argue that the skilled coordination of these musicians engaging in joint musical action is grounded in (implicit or explicit) knowledge of sensorimotor contingencies supporting music perception and production. These shared sensorimotor contingencies enable an interacting multi-agent group of musicians to co-create a shared, acoustic social world—forming a single complex system—as the interdependent actions of individual musicians give rise to emergent dynamics of an interacting musical ensemble.

Enactive sensorimotor contingency theory is a theory of perception, which describes perception as a process which is guided by action, emphasizing a “pre-conceptual, pre-linguistic form of understanding related to bodily and motor expertise” (Matyja & Schiavio 2013). Developed originally as an explanation for visual perception (O’Regan & Noë, 2001; Noë, 2004), the classic example of sensorimotor contingency theory is ‘seeing the whole tomato’. When we see a tomato, we don’t just see a two-dimensional gradient of colors and edges, but in some sense, we see the ‘whole’ tomato. Our awareness of the back of the tomato arises from our bodily knowledge of a repertoire of motor actions, the sensorimotor contingencies (SMCs), necessary in perceiving tomatoes: we know that if we were to perform a certain action (turn the tomato around), that we would see the back of a tomato. Enactivist accounts of music cognition place the perception of music in (implicit or explicit) bodily knowledge of a repertoire of motor actions and their effect on associated sensory stimulation, or knowledge of sensorimotor contingencies (Matyja & Schiavio, 2013). Rather than passive listeners, simply receiving a barrage of acoustic stimuli and later appraising it as musical (as in a traditional, cognitivist account of music perception), we perceive music through skilled action (Maes et al., 2014; Maes, 2016). The music listener learns to ‘manipulate’ the barrage of acoustic stimuli she hears through active (attentive) listening and skillful engagement with the musical environment (Krueger, 2009; Krueger, 2013). Musical training or experience enhances her knowledge of the sensorimotor contingencies involved in producing music, which enables her to selectively attend to increasingly more salient musical features when perceiving music. Knowledge of SMCs involves not only knowledge of what sound can be heard given a certain action, but also what action most likely caused or will cause a certain sound. This bidirectional association between action and perception enables individuals to plan and respond to their own action, and also to predict and coordinate with others through joint musical action—as listeners and players at the same time.

A series of studies by Drost et al. (2005, 2007) demonstrate that individuals with musical training were more susceptible to making mistakes in a forced production task due to incongruencies between visually and auditorily presented chords, and that the effect was stronger when the auditory stimulus presented was in the timbre of their own instrument. These studies indicate that musical training leads to more precise sensorimotor representations of the action necessary to produce a sensory stimulus (the heard chord). A number of piano timing experiments demonstrated that pianists (ignorant of the task condition, and told that they were performing with a live partner) were better able to play in time with recordings of themselves than of other musicians (Keller et al. 2007) or with others who were matched in terms of preferred performance tempo (Loehr et al. 2011). Each of these studies indicate that higher knowledge of SMCs enhances the participant's ability to be in time in a music production task by enhancing the participant's ability to plan and coordinate with a partner in joint musical activity.[2]

Rhythmic interaction in naturalistic music making, such as musical ensemble performance, relies on extending these simple sensorimotor contingencies—knowledge of the sound your instrument will make when you perform an action; knowledge of the sounds you'll hear based on the preferred tempo at which you play music—to more complex SMCs which take

---

[2] The original authors interpretation of these experiments and results was taken as support for the role of cognitive representations of the actions of self and other in musical activity. Thus, by their account, musicians are cognitively simulating the movement and production of the other musician. Under moderate accounts of SMC, these representations may be thought to include knowledge of SMCs, however under more radical accounts of SMC (with no role for representations) these studies may be taken to support the role of bodily knowledge/memory of the SMCs involved in music production and response.

into account the dynamics of two or more interacting agents, such as knowing what musical phrase you'll hear from your band mates in a jam session after you've each taken a certain set of musical actions. Humans excel at the precise timing and coordination of motor and acoustic output from multiple musicians in part because they excel at a skill called entrainment, which is where we are headed next.

Successful coordination within and between human individuals in music making may reflect a greater (implicit or explicit) knowledge of the sensorimotor contingencies involved in perceiving and producing musical events. Take our pianists for example, they are better able to synchronize with recordings of themselves (Keller et al. 2007) or with others who are matched in preferred performance tempo (Loehr et al. 2011). This is because the pianist (unknowingly) playing with a recording of herself has a very strong implicit knowledge of what actions it would take for her to produce the sound she hears from the recording. This strong knowledge of SMCs makes it easier for her to predict when and what she will hear, and enhances the strength of entrainment between musician and recording. It is thus easier for the pianist to form an interpersonal synergy alongside their own pre-recorded musical activity. The pianist who is playing with another who prefers similar tempi has a similarly strong knowledge of SMCs involved in producing the sounds they hear at their preferred tempo, enhancing the strength of entrainment at that tempo, and enabling the emergence of a tight interpersonal synergy.

Entrainment, or a specific form of coordination referring roughly to the ability to synchronize or to be together in time with one or more individuals, has been taken to "relate phenomenologically to a sense of social belonging" and has been conceived as one explanation for group cohesion and bonding that emerges from joint activities such as music making (Clayton, Sager & Will 2005). Our human ability to entrain with others ranges from the subconscious synchrony of repetitive motions (i.e. happening to walk in step with another) to the synchronization of intentional temporal events such as synchronizing melody and harmony in joint musical interaction. In temporal rhythmic processing, it is the *interaction* of the body, brain, and environment which results in the emergent phenomenon of sensorimotor and neural entrainment (Ross & Balasubramaniam, 2014). Interpersonal synchronous movement between two or more individuals may be further linked through synchrony of neural oscillations across individuals (Novembre et al., 2017) and has been found to occur in naturalistic social interactions among affiliative partners (Kinreich et al., 2017). Even when referring in part to neural phenomena, such multi-scale and multi-level coordination patterns likely relate to our phenomenological experience of being part of an extended social and cultural environment. Kirchoff and Kiverstein describe this feeling as “phenomenal attunement— the feeling of being at home in a familiar culturally constructed environment” (Kirchoff & Kiverstein, 2020). The interactions of a musical ensemble, specifically when the ensemble is made up of a group of musicians who have engaged in repeated rehearsals and joint musical interactions together, provides an ideally structured social and musical environment for that ensemble to exhibit an extended cognition, if not an extended conscious mind (Spivey, 2020, Kirchoff & Kiverstein, 2020).

In perhaps a less enactivist light, shared predictive *models* of sensorimotor contingencies developed during and as a result of group music making may give rise to group identity in a similar fashion to the “predictive perception of sensorimotor contingencies” which are proposed to underlie a sense of self (Seth, 2014). While radical enactivism maintains a strictly antirepresentational character of the nature of cognitive processes (c.f. Hutto & Myin, 2012), we do not take a stance in this debate in this chapter. Rather, we argue that enactive SMC theory grounds and enhances aspects of coordination in joint musical action in terms of bodily and environmental states, regardless as to whether these states are represented in the brain as models or wholly constituted by the bodily/environmental states themselves.

Individuals engaged in joint music making often join together in larger groups than these piano duets, ranging from a four-person quartet to a large chorus, orchestra, or even a stadium full of concert goers singing along with their favorite band on stage. Nevertheless, the shared sensorimotor contingencies among multi-agent groups of interacting musicians enable them to co-create a shared, acoustic social world. In doing so, they form collective interpersonal synergies, allowing the interdependent actions of individual musicians to give rise to the emergent dynamics of an interacting musical ensemble. By examining these synergies in the context of 4E cognition (Newen, Bruin, & Gallagher, 2018), we can see them as emerging from groups of agents who are *embodied* and *enactive*, as well as *embedded* in an environment, thus making their cognition *extended* across many interacting elements. That is, when the people and their instruments are well coordinated by virtue of their shared and co-created acoustic social world, they form one complex system that, by itself, bears a substantial statistical similarity to a mind.

References

Atherton, Gray, Natalie Sebanz, and Liam Cross. "Imagine All The Synchrony: The effects of actual and imagined synchronous walking on attitudes towards marginalised groups." *PloS one* 14.5 (2019): e0216585.

Balconi, Michela, and Giulia Fronda. "The dialogue between two or more brains: the “hyperscanning” for organization." *Frontiers in Psychology* 11 (2020).

Clayton, Martin, Rebecca Sager, and Udo Will. "In Time with the Music: The Concept of Entrainment and Its Significance for Ethnomusicology." *ESEM Counterpoint* 1, no. 82 (2005).

De Jaegher, Hanne., Ezequiel Di Paulo, & Shaun Gallagher. “Can social interaction constitute social cognition?” *Trends in Cognitive Sciences.* 14 no. 10 (2010): 441-447.

Drost, Ulrich C., et al. "When hearing turns into playing: Movement induction by auditory stimuli in pianists." *The Quarterly Journal of Experimental Psychology Section A* 58.8 (2005): 1376-1389.

Drost, Ulrich C., Martina Rieger, and Wolfgang Prinz. "Instrument specificity in experienced musicians." *Quarterly Journal of Experimental Psychology* 60.4 (2007): 527-533.

Engel, Andreas K., Friston, Karl J., & Kragic, Danica. (eds.), *The Pragmatic Turn. Toward Action-Oriented Views in Cognitive Science*, Cambridge MA-London, The MIT Press, (2016)

Hutto, Daniel D., and Erik Myin. *Radicalizing enactivism: Basic minds without content*. MIT Press, (2012).

Inagakuen Wind Orchestra. Welcome to the Imagination World (2009). Inagakuen Wind Orchestra. Youtube, uploaded by kamoshitamoyashi. 17 Feb. 2011, https://www.youtube.com/watch?v=-wJ9ZsgO3QI.

Keller, Peter E., Günther Knoblich, and Bruno H. Repp. "Pianists duet better when they play with themselves: on the possible role of action simulation in synchronization." *Consciousness and cognition* 16.1 (2007): 102-111.

Kinreich, Sivan, et al. "Brain-to-brain synchrony during naturalistic social interactions." *Scientific reports* 7.1 (2017): 1-12.

Kirchhoff, Michael D., and Julian Kiverstein. "Attuning to the World: The Diachronic Constitution of the Extended Conscious Mind." *Frontiers in Psychology* 11 (2020).

Krueger, Joel. "Enacting musical experience." *Journal of Consciousness Studies* 16.2-3 (2009): 98-123.

Krueger, Joel. "Empathy, enaction, and shared musical experience: evidence from infant cognition." The Emotional Power of Music: Multidisciplinary perspectives on musical arousal, expression, and social control. (2013): Oxford University Press, 26.

Loehr, Janeen D., Edward W. Large, and Caroline Palmer. "Temporal coordination and adaptation to rate change in music performance." *Journal of Experimental Psychology: Human Perception and Performance* 37.4 (2011): 1292.

Maes, Pieter-Jan, et al. "Action-based effects on music perception." *Frontiers in Psychology* 4 (2014): 1008.

Maes, Pieter-Jan. "Sensorimotor grounding of musical embodiment and the role of prediction: A review." *Frontiers in Psychology* 7 (2016): 308.

Matyja, Jakub, and Andrea Schiavio. "Enactive Music Cognition: Background and Research Themes." *Constructivist Foundations*, (2013): 351-57.

Misaki, Masaya, et al. "Beyond synchrony: the capacity of fMRI hyperscanning for the study of human social interaction." *Social Cognitive and Affective Neuroscience* 16.1-2 (2021): 84-92.

Newen, Albert, Leon De Bruin, and Shaun Gallagher, eds. *The Oxford Handbook of 4E Cognition*. Oxford University Press, (2018).

Noë, Alva. *Action in Perception*. Cambridge, MA: MIT Press, (2004).

Novembre, Giacomo, et al. "Interpersonal synchrony enhanced through 20 Hz phase-coupled dual brain stimulation." *Social Cognitive and Affective Neuroscience* 12.4 (2017): 662-670.

O'Regan, J. Kevin, and Alva Noë. "A Sensorimotor Account of Vision and Visual Consciousness." *Behavioral and Brain Sciences* 24 (2001): 939-1031.

Proksch, Shannon, et al. "Coordination Dynamics of Multi-agent Human Interaction." PsyArXiv, 25 Jan. 2021.

Riley, Michael A., et al. "Interpersonal synergies." *Frontiers in psychology* 2 (2011): 38.

Ross, J. M., & Balasubramaniam, R. "Physical and neural entrainment to rhythm: Human sensorimotor coordination across tasks and effector systems." *Frontiers in Human Neuroscience* 8 (2014): Article 576. doi: 10.3389/fnhum.2014.00576

Schirmer, Annett, Merle Fairhurst, and Stefanie Hoehl. "Being 'in sync'—is interactional synchrony the key to understanding the social brain?" *Social Cognitive and Affective Neuroscience* 16.1-2 (2021): 1-4.

Seth, Anil K. “A predictive processing theory of sensorimotor contingencies: Explaining the puzzle of perceptual presence and its absence in synesthesia.” *Cognitive neuroscience* vol. 5,2 (2014): 97-118. doi:10.1080/17588928.2013.877880

Shimizu, Daisuke. Welcome to the Imagination World (2016). www.windrep.org/Welcome_to_the_Imagination_World.

Shockley, Kevin, Marie-Vee Santana, and Carol A. Fowler. "Mutual interpersonal postural constraints are involved in cooperative conversation." *Journal of Experimental Psychology: Human Perception and Performance* 29.2 (2003): 326.

Spivey, M. *Who You Are: the Science of Connectedness*. MIT Press, (2020).

Walton, Ashley E., et al. "Creating time: Social collaboration in music improvisation." *Topics in cognitive science* 10.1 (2018): 95-119.