FACULTEIT LETTEREN EN WIJSBEGEERTE

UNIVERSITEIT GENT

Université de Lille

Université de Lille
Ghent University
Ecole Doctorale Science de l'Homme et de la Société
Laboratoire PSITEC - ULR 4072 Psychologie : Interactions, Temps, Emotions, Cognition

Human rhythmic interactions

Coordination dynamics and informational coupling

Mattia Rosso

Dissertation submitted to obtain the degrees of Doctor of Art Sciences (University of Ghent, BE) and Doctor of Psychology (University of Lille, FR)

Public defense on september 21th, 2023

Supervisors Prof. Marc Leman, *Ghent University, Faculty of Arts and Philosophy, IPEM Institute for Psychoacoustics and electronic music*

Prof. Pieter-Jan Maes, *Ghent University, Faculty of Arts and Philosophy, IPEM Institute for Psychoacoustics and electronic music*

Prof. Séverine Samson, *Université de Lille, Équipe Neuropsychologie et Audition, Laboratoire PSITEC – ULR 4072*

Joint examination board

Prof. Peter Keller (*Aarhus University, Department of Clinical Medicine*), Reviewer

Prof. Benjamin Morillon (*University of Aix-Marseille, Faculty of Medicine*), Examiner

Prof. Sonja Kotz (*Maastricht University, Department of Neuropsychology & Psychopharmacology*), *Examiner*

Dr. Lise Hobeika (*University of Lille, Department of Psychology*), Examiner

Dr. Joren Six (*Ghent University, Faculty of Arts and Philosophy*), Examiner

Président du jury: Prof. Sylvie Nozaradan (*Université Catholique de Louvain, Faculty of Medicine and Dentistry*), Reviewer

FACULTEIT LETTEREN
EN WIJSBEGEERTE

Human rhythmic interactions

Coordination dynamics and informational coupling

Mattia Rosso

Dissertation submitted to obtain the degrees of Doctor of Art Sciences (UGent, BE) and
Doctor of Psychology (ULille, FR)
Academic year 2022-2023

UNIVERSITEIT
GENT

Université
de Lille

Supervisors

Prof. Marc Leman

Ghent University, Faculty of Arts and Philosophy

IPEM Institute for Psychoacoustics and electronic music

Prof. Pieter-Jan Maes

Ghent University, Faculty of Arts and Philosophy

IPEM Institute for Psychoacoustics and electronic music

Prof. Séverine Samson

Université de Lille, Équipe Neuropsychologie et Audition,

Laboratoire PSITEC – ULR 4072

Joint examination board

Prof. Peter Keller (*Aarhus University, Department of Clinical Medicine*)

Prof. Benjamin Morillon (*University of Aix-Marseille, Faculty of Medicine*)

Prof. Sonja Kotz (*Maastricht University, Department of Neuropsychology & Psychopharmacology*)

Dr. Lise Hobeika (*University of Lille, Department of Psychology*)

Dr. Joren Six (*Ghent University, Faculty of Arts and Philosophy*)

Président du jury: Prof. Sylvie Nozaradan (*Université Catholique de Louvain, Faculty of Medicine and Dentistry*)

Acknowledgements

Before Professor Marc Leman became my supervisor, I had read his book four times to wrap my head around it. While much of its content was beyond my comprehension at the time, two things became crystal clear from the very first read. The first was that there was something special hidden in the chapter titled *Entrainment*. I sensed something awaiting my exploration, and I'm profoundly grateful for the five years I was given to delve into that. The second revelation was in the acknowledgments of his book: somehow, Marc had assembled a 'dream team'. This made me realize, both then and now, that I aim to form my own dream team someday. While my personal contributions to research might just be a drop in the ocean, I am at peace with it. My true ambition is to inspire the talented researchers of the future, to guide them and, in the bigger scheme of things, to keep the wheel of science spinning. Just as you have been doing for the last many years, Marc. Thank you for that. Thank you for giving me all the freedom and trust I needed, and for being a mentor.

Marc warned me the project was ambitious, that it would be challenging. So, I packed my stuff and moved to Ghent with big dreams and a small piece of paper, where I had written: ‘talk to Maes’. During my first lunch at IPEM, I sat at the table next to a cheerful man, who seemed close to my age. A very friendly guy indeed. “Any chance that you know Professor Pieter-Jan Maes?”, I asked. It turned out that I was speaking to the very person I was seeking. I am glad I did, because that encounter ignited the chain of events leading to the completion of this book. I am fortunate that you became my co-supervisor, Pieter-Jan, and I eagerly anticipate the discussions and explorations our future holds.

Next to the opportunities I was given at IPEM, I am extremely grateful to Professor Séverine Samson for her willingness to co-supervise my work in the pursuit of a joint-PhD. I say this even more heartfeltly for being so supportive beyond the difficulties we had to deal with. I am glad I had the chance to work with you, Séverine, and I truly hope that my work can contribute in some way to your research, because I believe it is crucial for humankind. Thank you for opening the door of your home to me, both metaphorically and literally. Feeling part of your extended family together with Andres and Lise was special.

Even though his name is not listed in my doctoral committee, I must thank Professor Peter Vuust for his warm welcome at the Center for Music in the Brain in Århus, time and again. I've always felt a sense of belonging there, and I am thrilled at the idea that this feeling will continue to grow. Leonardo, Ole, Sander, David, I've learned so much from you, and I'm grateful that you have been part of my journey.

I have the privilege of saying this book was vetted by the best jury I could hope for: Professors Peter Keller, Sylvie Nozaradan, Benjamin Morillon, and Sonja Kotz; along with Doctors Lise Hobeika and Joren Six. We all have some aspirations, but no one aspires in the dark. To all of you, jury members, know that you have been among the lights guiding me in the pursuit of my project. Thank you for your work and for your precious time.

Moving on to Marc's dream team, huge thanks to IPEM. Having joined at such a dynamic period with so many amazing folks coming and going, it is impossible to give everyone the shout-out they deserve here. But special mentions to:

Andrea, for being my first friend in a foreign land and the producer of the first version of the *drifting metronomes*. There has never been a second version.

Bavo, for our metastability and our cooperation/competition balance. More about this throughout the book.

Lousin, for pushing each other to our limits when we needed it, and for slowing each other down when we needed it more.

Marc V, for how you navigate emotions and mathematics.

Joren, for your stoic presence, for being there for us like a precious rock.

Ola, for a bond that I would not know how to call, if not: music.

Canan, for being a treasure.

This book would not have been possible without all the volunteers who took part in the experiments. Your brainwaves and countless finger taps have been invaluable. A special thanks to those whose enthusiasm turned tedious tasks into exciting learning experiences, and to the unique few who became good friends of mine. Kudos to Bavo and Lousin, for proof-reading the final draft of this book and for sharing your insights.

Now comes the most cherished part. As nomadic as I may be, I didn't fully grasp how big of a step it was to settle in this country for such a long time. Outside of work, I thank all the friends and special people who were with me between my peaks and troughs. In case you are reading this, you know who you are.

Finally, to my family:

Rachele, I hope that something in me can be of inspiration for you, whatever your path may be. Make it a unique one.

Emma, maybe one day you'll pick up this book and get a glimpse of what your big brother was doing all those miles away while you were growing up.

Mum, I know that whatever my path may be, your unconditional love wouldn't change by an inch. That is why I know I can go anywhere, I can do anything.

Dad, most of all, this book is dedicated to you, my example of perseverance. For countless evenings have I peeked through the keyhole, watching you sit at your desk, engrossed in finishing your book. Now I know how it feels to sit on the other side of that door.

Table of Contents

English summary

This dissertation investigates rhythm as a foundational element of human functioning, and a substratum for interactions between individuals. Music serves as an exemplary demonstration of this dimension. In particular, ensemble performances showcase precise and flexible temporal coordination, resulting from the interplay between perceptual and motor processes. These dynamic interactions lead to the emergence of organic, structured, and dynamic musical performances, encompassing the very essence of rhythm's role in interpersonal coordination. However, the purpose of the dissertation extends beyond music, aiming to explore the dynamics underpinning interpersonal coordination and the organization of individual rhythms into collective behavior. These organizational principles are generalizable to all human interactions.

Rooted in key theoretical concepts, such as the central role of perception in mediating the coupling between individuals and enabling their coordination, the dissertation unfolds within an overarching methodological framework. Novel experimental paradigms and signal processing techniques are developed with the goal of eliciting and quantifying the dynamics underlying dyadic behavior. From this approach, the multifaceted nature of perceptual coupling is illustrated, and its impact on coordination dynamics between individuals is investigated in depth. Throughout a series of studies, fundamental dimensions of informational coupling via sensory channels are identified and manipulated to assess their specific effect on the interaction. Starting from the direct comparison of visual and auditory couplings (Chapter 2), the investigation expands to subsidiary dimensions, such as their differential access to kinematic information (Chapter 3) and the role of perspective taking (Chapter 4).

Alongside the assessment of dyadic behavior, the dissertation also delves into the neural dynamics underlying these interactions. By analyzing electroencephalographic (EEG) data recorded simultaneously from interacting participants, the interplay between behavioral and neural oscillators is investigated as a potential mechanism underpinning interpersonal synchronization (Chapter 5). To enhance this investigation, computer-generated auditory stimuli are utilized to induce and enable the measurement of neural entrainment, defined as the alignment of endogenous brain oscillations to environmental rhythms (Chapters 6 and 7).

Finally, building on the methodologies developed throughout the dissertation, an examination of rhythmic musical interactions in aging individuals with varying degrees of neurocognitive disorder is conducted. Two datasets collected by the PSITEC Laboratory

(University of Lille, FR), one involving a dyadic rhythmic task with a musical therapist, and another one featuring a synchronization task to isochronous and perturbated auditory stimuli, are analyzed. This contribution showcases how to overcome challenges related to data acquisition with this vulnerable population, thereby allowing maximal inference on their sensorimotor synchronization abilities (Chapter 8).

The conclusion of this dissertation is the integration of concepts and methodologies that offer a comprehensive understanding of the role of rhythm in human interactions. The dissertation maintains a strong systemic perspective, considering the dyad as a coupled collective unit. It enriches theoretical and methodological dimensions across disciplines, paving the way for future research directions at the intersection of behavioral neuroscience, systematic musicology, and psychology.

Nederlandstalige samenvatting

Dit proefschrift onderzoekt ritme als een fundamenteel element van menselijk functioneren en als substraat voor interacties tussen individuen. Muziek is hier een uitstekende demonstratie van. Met name ensemble-uitvoeringen laten een nauwkeurige en flexibele temporele coördinatie zien als resultaat van de wisselwerking tussen perceptuele en motorische processen. Deze dynamische interacties resulteren in het ontstaan van een organisch, gestructureerde, dynamische muzikale uitvoering, die de essentie van de rol van ritme in interpersoonlijke coördinatie tonen. Het doel van het proefschrift reikt echter verder dan muziek. Het doel van dit proefschrift is om de dynamiek die ten grondslag ligt aan interpersoonlijke coördinatie en de organisatie van individuele ritmes in collectief gedrag te onderzoeken. Dit is te generaliseren naar alle menselijke interacties.

De dissertatie baseert zich op belangrijke theoretische concepten, zoals de centrale rol van perceptie in het bemiddelen van de koppeling tussen individuen en hun coördinatie, die binnen een overkoepelend methodologisch kader beschreven worden. Nieuwe experimentele paradigma's en signaalverwerkingstechnieken worden gepresenteerd die ontwikkeld zijn om de dynamiek van attractoren in dyadisch gedrag uit te lokken en te kwantificeren. Vanuit dit perspectief wordt de veelzijdige aard van perceptuele koppeling belicht en wordt de invloed ervan op de coördinatiedynamiek tussen individuen onderzocht. In een reeks studies worden verschillende fundamentele dimensies van informatie-uitwisseling via sensorische kanalen geïdentificeerd en gemanipuleerd om hun specifieke effect op de interactie te beoordelen. Beginnend met de directe vergelijking van visuele en auditieve koppelingen (hoofdstuk 2), breidt het onderzoek zich uit naar enkele aangrenzende dimensies, zoals de verschillende manieren van toegang tot kinematische informatie (hoofdstuk 3) en de rol van perspectief in een omgeving gemedieerd door virtual reality technologie (hoofdstuk 4).

De neurale dynamiek onderliggend aan dyadisch gedrag wordt ook onderzocht. Door elektro-encefalografische (EEG) gegevens te analyseren die gelijktijdig bij beide participanten werd opgenomen terwijl zij met elkaar communiceren, wordt de wisselwerking tussen gedrags- en neurale oscillatoren onderzocht als een potentieel mechanisme dat ten grondslag ligt aan interpersoonlijke synchronisatie (Hoofdstuk 5). Om dit onderzoek te versterken worden computer gegenereerde auditieve stimuli gebruikt om *neural entrainment* te induceren en te meten, gedefinieerd als het aligneren van endogene hersenoscillaties met omgevingsritmes (hoofdstuk 6 en 7).

Tenslotte wordt, voortbouwend op de in het proefschrift ontwikkelde methoden, een onderzoek uitgevoerd naar ritmische muzikale interacties bij ouder wordende personen met verschillende niveaus van neurocognitieve stoornissen. Twee datasets die verzameld zijn door het PSITEC laboratorium (Universiteit van Lille, FR) worden geanalyseerd: één met een dyadische ritmische taak met een muzikale therapeut en een andere met een synchronisatietaak op isochrone en verstoorde auditieve stimuli. Deze bijdrage laat zien hoe men uitdagingen in verband met dataverzameling bij deze kwetsbare populatie kan overwinnen, waardoor de conclusies over hun sensomotorische synchronisatiecapaciteiten worden gemaximaliseerd (Hoofdstuk 8).

De culminatie van dit proefschrift bevindt zich in de integratie van concepten en methodologieën ten behoeve van een uitgebreider begrip omtrent de rol van ritme in menselijke interacties. Het heeft een sterk systemisch perspectief, waarbij het duo wordt beschouwd als een gekoppelde collectieve eenheid. Als zodanig verrijkt het zowel de theoretische als de methodologische dimensies van dit veld en baant het de weg voor toekomstige onderzoeksrichtingen in de neurowetenschappen, systematische musicologie, en psychologie.

Résumé de la thèse

Cette dissertation examine le rythme comme un élément fondamental du fonctionnement humain et un substrat pour les interactions entre individus. La musique sert comme démonstration exemplaire de cette dimension. En particulier, les performances en ensemble mettent en évidence une coordination temporelle précise et flexible, résultant de l'interaction entre les processus perceptifs et moteurs. Ces interactions dynamiques conduisent à l'émergence de performances musicales organiques, structurées et dynamiques, englobant l'essence même du rôle du rythme dans la coordination interpersonnelle. Cependant, le but de la dissertation va au-delà de la musique, visant à explorer la dynamique sous-tendant la coordination interpersonnelle et l'organisation des rythmes individuels en comportement collectif, en principe généralisable à toutes les interactions humaines.

Enracinée dans des concepts théoriques clés, tels que le rôle central de la perception dans la médiation du couplage entre individus et leur coordination, la dissertation se déploie dans un cadre méthodologique global. De nouveaux paradigmes expérimentaux et des techniques de traitement du signal sont développés dans le but de susciter et de quantifier la dynamique sous-jacente au comportement dyadique. De ce point de vue, la nature multifacette du couplage perceptif est mise en évidence, et son impact sur la dynamique de coordination entre individus est étudié en profondeur. Tout au long d'une série d'études, différentes dimensions fondamentales du couplage informationnel par les canaux sensoriels sont identifiées et manipulées pour évaluer leur effet spécifique sur l'interaction. En partant de la comparaison directe des couplages visuels et auditifs (Chapitre 2), l'investigation s'étend à certaines dimensions subsidiaires, telles que leur accès différentiel aux informations cinématiques (Chapitre 3), et le rôle de la prise de perspective (Chapitre 4).

Parallèlement à l'évaluation du comportement dyadique, la dissertation se penche également sur la dynamique neuronale sous-jacente à ces interactions. En analysant les données électroencéphalographiques (EEG) enregistrées simultanément chez des participants interagissant, l'interaction entre les oscillateurs comportementaux et neuronaux est étudiée comme un mécanisme potentiel sous-tendant la synchronisation interpersonnelle (Chapitre 5). Pour améliorer cette investigation, des stimuli auditifs sont utilisés pour induire et mesurer l'entrainement neural, à savoir l'alignement des oscillations cérébrales endogènes sur les rythmes environnementaux (Chapitres 6 et 7).

En fin, en s'appuyant sur les méthodologies développées tout au long de la thèse, une examination des interactions musicales rythmiques est menée chez des individus âgés présentant divers niveaux de troubles neurocognitifs. Deux jeux de données collectés par le laboratoire PSITEC (Université de Lille, FR) sont analysés : l'un impliquant une tâche rythmique dyadique avec un musicothérapeute, et un autre présentant une tâche de synchronisation à des stimuli auditifs isochrones et perturbés. Cette contribution met en évidence la surmonte des défis liés à l'acquisition de données avec cette population vulnérable, maximisant ainsi les inférences sur leurs capacités de synchronisation sensorimotrice (Chapitre 8).

L'apogée de cette thèse est l'intégration de concepts et de méthodologies pour offrir une compréhension complète du rôle du rythme dans les interactions humaines. Elle maintient une perspective systémique forte, considérant la dyade comme une unité collective couplée. En tant que tel, il enrichit à la fois les dimensions théoriques et méthodologiques à travers les disciplines, ouvrant la voie à de futures directions de recherche à l'intersection des neurosciences comportementales, de la musicologie systématique, et de la psychologie.

Research output (2019-2023)

Scientific publications

A1 peer-reviewed publications

Total of 7 articles; 4 as first author; 3 as second author, H-index 4.

*** Rosso, M.**, Moens, B., Leman, M., & Moumdjian, L. (2023). Neural entrainment underpins sensorimotor synchronization to dynamic rhythmic stimuli. *NeuroImage*, 120226. [Impact Factor 7.4. Q1 Cognitive Neuroscience (5/107); Q1 Neuroimaging (2/14)].

*** Rosso, M.**, Heggli, O. A., Maes, P. J., Vuust, P., & Leman, M. (2022). Mutual beta power modulation in dyadic entrainment. *NeuroImage*, 257, 119326. [Impact Factor 7.4. Q1 Cognitive Neuroscience (5/107); Q1 Neuroimaging (2/14)].

*** Rosso, M.**, Maes, P. J., & Leman, M. (2021). Modality-specific attractor dynamics in dyadic entrainment. *Scientific Reports*, 11(1), 1-13. [Impact Factor 4.997. Q1 Multidisciplinary (11/120)].

*** Rosso, M.**, Leman, M., & Moumdjian, L. (2021). Neural entrainment meets behavior: the stability index as a neural outcome measure of auditory-motor coupling. *Frontiers in Human Neuroscience*, 15. [Impact Factor 3.473. Q2 Behavioral neuroscience (29/79)].

Vidal, M., **Rosso, M.**, & Aguilera, A. M. (2021). Bi-Smoothed Functional Independent Component Analysis for EEG Artifact Removal. *Mathematics*, 9(11), 1243. [Impact Factor 2.884. Q1 Mathematics (75/378)].

Dell'Anna, A., **Rosso, M.**, Bruno, V., Garbarini, F., Leman, M., & Berti, A. (2021). Does musical interaction in a jazz duet modulate peripersonal space?. *Psychological Research*, 85(5), 2107-2118. [Impact Factor 2.956. Q1 Arts and Humanities (43/306), Q2 Experimental and Cognitive Psychology (44/148)].

Van Kerrebroeck, B., **Rosso, M.**, & Maes, P. J. (2020). Linking embodied coordination dynamics and subjective experiences in musical interactions: a renewed methodological paradigm. DOCUMENTA, 38(1), 38-60.

A1 peer-reviewed papers under review

Total of 1 article; 1 as first author

*** Rosso, M.**, van Kerrebroeck, B., Maes, P. J., & Leman, M. (*accepted for publication*). Embodied perspective taking enhances interpersonal synchronization. *iScience*. [Impact Factor 6.1. Q1 Multidisciplinary].

A2 peer-reviewed publications

Total of 1 article; 1 as second author

Moumdjian, L., **Rosso, M.**, Moens, B., De Weerdt, N., Leman, M., & Feys, P. (2022). A case-study of a person with multiple sclerosis and cerebellar ataxia synchronizing finger-taps and foot-steps to music and metronomes. *Neuroimmunology Reports*, 2, 100101.

***** *These scientific papers are included in this dissertation, either in their published version or incorporating the most recent revisions.*

Scientific conferences

Rosso, M., Maes P., Leman. "Dyadic rhythmic interactions. Coordination dynamics and informational coupling".

The 17th International Conference on Music Perception and Cognition, August 24th-28th 2023 – Tokyo (JP)

(Oral presentation)

Rosso, M., Leman, M., Moumdjian, L. "Event-related frequency adjustment. A methodology for investigating neural entrainment".

Sysmus, September 7-9th 2022 – Ghent (BE)

(Oral presentation)

Rosso, M., Leman, M., Moumdjian, L. "Neural entrainment meets behaviour: the Stability Index as a neural outcome measure of auditory-motor coupling".

Rhythm production and perception workshop, June 22-25th 2021 – Oslo (NO)

(Oral presentation)

Rosso M., Heggli O.A., Maes P., Vuust P., Leman M. "Drifting metronomes. Cooperation and competition in dyadic entrainment".

Neuromusic VII, June 18-21th 2021 – Aarhus (DK)

(Poster presentation)

Rosso M. "Interactive settings for interactive brains".

ANT Neuromeeting, January 15-18th 2020 – Beaune (FR)

(Poster presentation)

Prologue

This work is about human interaction, and the rhythms underneath.

Think of the myriad of ways people communicate, move, engage in sports, play, work, perform, cook, love, and fight with one another. It's astounding how naturally they harmonize their movements over time. By gazing through the vast plethora of human action, one can see there is a rhythm to it, and that it is organized in emerging patterns.

An underlying force appears to weave through our lives, pulling us together and pushing us apart in an oscillating dynamic between unity and individuality. It is as though we are one, then two, then one again, ceaselessly drawn into the ebb and flow of connection and separation. We cooperate, we compete, then we cooperate once again. Friendships bloom and wither, lovers navigate an endless cycle of attachment and detachment. How can incongruous desires for closeness and independence coexist within couples, as we perpetually strive for balance amidst turbulent fluctuations? It is almost as if we were magnets, attracting and repelling one another in a constant interplay of forces.

Perhaps these metaphors evoke images that resonate with our lived experiences. And whilst that all may sound poetic, this work strives for scientific understanding. Both the metaphor in the poet's eyes and the model in the scientist's mind aim at capturing the essence of our complex shared experiences, like getting to the core of this magnetic dance. What is a metaphor if not a model, after all? What is a metaphor, if not a means of illustrating qualities that underlie the physical manifestations of a phenomenon? In scientific terms, we would claim that the interaction can be explained by a set of laws, which are there for the scientist to discover.

This dissertation aims to explore human interactions at their most fundamental level. By peeling the outer layers of complex interactions, I aim to get at the lowest common denominator of the manifold manifestations of our relational dynamic. I will therefore focus on the rhythmic aspects, seeking to uncover the underlying principles and patterns that govern basic interactions with scientific rigor, while trying not to distort their essence. To achieve this, I will examine the conditions under which interpersonal interactions occur, as well as their dynamics and their underlying control mechanisms as they unfold over time. In this endeavor, I will investigate the neural underpinnings of rhythmic behavior, and try to bridge them to major theories of brain functioning in interaction with the environment and other human actors. Moving on this fundamental level, I strive to unravel the complex tapestry of human connections, while acknowledging that the beauty

and richness of our interactions will never be encapsulated by a single metaphor nor by a model. Yet, we must accept that scientific endeavors come with the inevitable simplification of complex phenomena.

After turning this page, the reader will pardon the temporary suspension of the poetic, for the sake of scientific writing. I hope you can still find some beauty in it, or at least some new insights. In any case, I hope you will enjoy the reading.

1

Theoretical and methodological framework

Major theoretical concepts

The purpose of this dissertation is to unravel the dynamics that are at play when two humans coordinate with each other, and that allow their individual rhythms to give rise to organized forms of collective behavior. In this endeavor, the dissertation will balance its focus between individual and dyadic levels of analysis. Namely, it will be necessary to define rhythm and how a human interacts with it as a single organism, before investigating it as a mediator and substratum for interpersonal interactions.

In what follows, we navigate major theoretical concepts, smoothly transitioning in and out of the social domain to provide a cohesive presentation of the topic of human rhythmic interactions. After illustrating the fundamentals, we will discuss how these relate to major theories dominating current debates in cognitive and social neurosciences. The theoretical introduction will be followed by the key methodological concepts that grounded and oriented the program of empirical research throughout this PhD project. In conclusion of this introductory chapter, as we lay out the structure of the dissertation in the chapters overview, we will present the research questions that will be addressed by each individual study.

Rhythm

At the backbone of human movement and interactions lies rhythm (P. E. Keller, Novembre, & Hove, 2014), the foundational element that structures events in time, a crucial dimension of action, life, and nature itself. Humankind has achieved a sublime feat in crafting the most sophisticated expression of temporal organization in the cultural artifact we call 'music,' artistically illustrating the progression and evolution of these sequences within compositions (Alperson, 1980). In this context, rhythmic patterns can be defined as recurrent sequences exhibiting temporal hierarchies (Lenc et al., 2021), with varying degrees of complexity and predictability in relation to an underlying regular pulse (Vuust, Heggli, Friston, & Kringelbach, 2022). Notably, the presence of music as a conserved cultural practice has been observed across all cultures throughout human history, globally, with rhythm as its fundamental element (Jacoby & McDermott, 2017; Polak, Jacoby, & Fischinger, 2018).

Besides the evidence that humans are biologically predisposed to perceive and produce musical rhythms (Cannon & Patel, 2021; Lenc et al., 2021; Merchant, Grahn, Trainor, Rohrmeier, & Fitch, 2015; Niarchou et al., 2022), there is an inherently social aspect to this

natural inclination (P. E. Keller et al., 2014). Music making is a lively interactive practice, manifested in the traditional forms of solo or ensemble playing, dancing, shared listening, and in novel forms allowed by the technological development of extended musical spaces (Van Kerrebroeck, 2023). The rhythm in music offers humans an affordance for embodied interactions, engaging them in a process of alignment between body kinematics and acoustic structures (Marc Leman, 2016). The mapping between bodily (motoric) and sonic activation provides the basis for pragmatic communication among individuals via the encoding and decoding of intentional expressive gesturing (M. Leman & Godøy, 2010), and a rich substrate for cohesive group interactions. This is the reason why joint music making induces pro-social effects on individuals (Savage et al., 2020), whether they sing, dance, or attend concerts together (Onderdijk, 2022). Regardless of the particular musical context of activity, rhythm stands out as a key element for bonding individuals (Hove & Risen, 2009; Marsh, Richardson, & Schmidt, 2009), thanks to a positive feedback loop between synchronization and socio-cognitive skills mediated by empathy (for a review, see Tzanaki, 2022).

By its nature, music provides the most articulated expression of meaningful temporal structuring, and can thus be considered the art of time at its finest (Alperson, 1980). Tracing back from this sophisticated human practice, we can observe simpler instances of temporal organization in action. It becomes evident that rhythm characterizes various other aspects of social life, such as verbal (Cummins, 2009b; Grabe & Low, 2013; Ramus, Nespor, & Mehler, 2000) and non-verbal communicative transactions (Greenfield, 1994a; Merker, Madison, & Eckerdal, 2009), as well as basic individual behaviors like walking (Styns, van Noorden, Moelants, & Leman, 2007).

The normal execution of rhythmic behaviors is essential in daily life. This becomes evident when neuropathology impairs the neuroanatomical structures required for this functioning, heavily compromising quality of life (Grabli et al., 2012; Moumddjian, 2020). In these instances, leveraging the common rhythmic substratum between the movement in the body and the movement in the music has opened a promising avenue for music-based interventions in neurorehabilitation (Bella, Benoit, Farrugia, Keller, & Obrig, 2017; Dotov, de Cock, Geny, Driss, & Garrigue, 2017; Koshimori & Thaut, 2018; Moumdjian, Moens, Maes, Van Nieuwenhoven, et al., 2019; Nombela et al., 2013) and well-being (Bardy, Hoffmann, Moens, Leman, & Dalla Bella, 2015; Buhmann, Moens, Van Dyck, Dotov, & Leman, 2018; Micheline Lesaffre, 2018; B. Moens, 2018; Van Dyck et al., 2015).

Underneath the overt behavioral manifestations of bodily rhythms, zooming in on the underlying biological activity through the lens of electrophysiology, we can see that the activation of both central and autonomic nervous systems is inherently rhythmic. Heart

activity, respiration, digestion, and brain activity all exhibit repeated patterns, undergoing cycles which continually influence the state of organisms (Bardy et al., 2015; Bron & Furness, 2009; Buzsáki, Logothetis, & Singer, 2013; Shaffer, McCraty, & Zerr, 2014; Van Kerrebroeck & Maes, 2021). A particular state will in turn modulate ecological behavior, as different phases result in alternating instances of high and low excitability, determining individuals' disposition to act, perceive, and evaluate their environment (Criscuolo, Schwartze, & Kotz, 2022). These rhythmic modulations unfold across multiple timescales, from monthly hormonal cycles (Menaker & Menaker, 1959; Raible, Takekata, & Tessmar-Raible, 2017), to daily circadian rhythms (Rivera & Huberman, 2020), multi-hour digestive processes (Bron & Furness, 2009), and sub-second fluctuations of attention span (Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2008) and perceptual cycles (Schroeder & Lakatos, 2009; Schroeder, Wilson, Radman, Scharfman, & Lakatos, 2010). As we continue this journey from macro to micro, we can still observe that rhythm is fundamental to the functioning of every subcomponent, even down to the observation that unique classes of neurons discharge action potentials with their own intrinsic dynamics (Kepecs & Fishell, 2014; Zeng & Sanes, 2017).

We posit that understanding ecological human behavior from a brain-body-environment perspective hinges on the mechanisms that underpin *interactions among rhythms*, across multiple timescales. These interactions serve two main purposes: 1) they maintain a dynamic internal organization within the organism, and 2) they allow for adaptation to the ever-changing external environment. Internal rhythms are organized in an architecture of functional dependencies, as revealed by patterns of co-variation among signals from different parts of the human body (Criscuolo et al., 2022). This structure provides a substrate for functional coupling within the body and the brain, enabling a complex interplay that maintains homeostatic equilibrium and directs behavior through state- or task-specific configurations.

Environmental rhythms, present in nature itself, occur at different timescales and levels of predictability. For instance, fluctuations of solar light and temperature throughout the day are highly predictable at different times of the year. In contrast, a piece of music can challenge the listener by varying rhythmic complexity from one moment to the next. (Matthews, Witek, Heggli, Penhune, & Vuust, 2019; Matthews, Witek, Lund, Vuust, & Penhune, 2020; Vuust, Dietz, Witek, & Kringelbach, 2018; Vuust & Witek, 2014; Witek, Clarke, Wallentin, Kringelbach, & Vuust, 2015). Social contexts add a layer of complexity, since predicting others' actions is hindered by their hidden intentional states and by the variability inherent to human movement (Stergiou & Decker, 2011). Despite their fundamental differences, these exemplary phenomena share a common feature: they tend to unfold in cycles, which interact with the cyclic processes internal to the perceiver.

For instance, consider our circadian rhythms that adapt to light among a variety of external time-keeping cues (Lazzerini Ospri, Prusky, & Hattar, 2017; Rivera & Huberman, 2020), or our predisposition to dance to music up to a certain level of predictability (Matthews et al., 2020; Stupacher, Hove, Novembre, Schütz-Bosbach, & Keller, 2013; Witek et al., 2015). Similarly, we can dance along with a partner in a silent process of mutual adaptation based on a shared common pulse (Lenc et al., 2021).

In summary, the internal rhythmic organization provides humans with a flexible temporal scaffolding that enables interactions with environmental rhythms of varying predictability. Distilling the complexity of interactions between rhythmic processes is the work of a whole dissertation, not merely of a single page. Nevertheless, it is important to already introduce the concepts of *entrainment* and *synchronization*, as they are arguably the most basic of these mechanisms. Strictly speaking, entrainment is defined as the one-way synchronization of an oscillating system to an external rhythmic driving force (Lakatos, Gross, & Thut, 2019). S*ynchronization*, a concept closely related to it, is the bidirectional process leading two systems to oscillate simultaneously or at regular temporal intervals (Pikovsky, Rosenblum, Kurths, & Synchronization, 2001). For example, in physics, when two oscillating pendula share a mechanical support, they synchronize through an adaptation process in which the swinging of one pendulum adjusts to match the rhythm of the other (Huygens, 1888). It is crucial to emphasize that entrainment is the process that leads to a synchronized state, and not the synchronized state itself. Throughout this work, we will maintain consistency in this terminology and adhere to the fundamental definitions of these physical phenomena.

Coordination

In order to understand the functioning of an organism from an ecological perspective, it is first necessary to distinguish the boundaries separating its internal organization from the external environment, and identify the mechanisms that enable interaction and navigation within environmental affordances. Across spatial and temporal scales, biological units form interconnected aggregates, each serving specific goals or functions. Neurons interact within networks, networks interact within the brain, the brain interacts with the entire body, and the body interacts with the environment (J. A. Scott Kelso, Dumas, & Tognoli, 2013). These sets of linked components or subsystems influence each other through spatial and temporal coordination, leading to the emergence of collective behaviors that serve specific biological functions (J. A. Scott Kelso, 2009).

Inspired by the physics of complex dynamical systems, the discipline of coordination dynamics approaches human behavior and cognition with a strong emphasis on the systemic properties emerging from the interaction between components. The seminal work of Scott Kelso (J. A. S. Kelso, 1995) was pivotal in laying the foundation for this discipline, focusing on the spatiotemporal relationships between units and elucidating the principles and control mechanisms underlying patterns of collective behavior. The relevance of this framework to this dissertation lies in its focus on interactions among rhythmic units at a higher level, where parsimonious descriptions of complexity generalize to human behavior. Whilst inter-limb rhythmic coordination was the 'workhorse' of the early studies in this tradition, a core tenet of the framework is that the same dynamic principles apply to the coordination among different parts of an organism, between organisms themselves, and between organisms and their environment (J. A. S. Kelso, 1995; Kugler & Turvey, 2015). In its quest for level-independent principles, coordination dynamics proposes that if the dynamics describing the coordination within an organism are sufficiently general, they should also encompass basic interactions between humans (Tognoli, Zhang, Fuchs, Beetle, & Kelso, 2020). This perspective allowed the field to extend into the social domain, scaling the analysis up to the interpersonal level (Guillaume Dumas, de Guzman, Tognoli, & Kelso, 2014; Guillaume Dumas, Kelso, & Nadel, 2014; J. A. Scott Kelso et al., 2013; R. C. Schmidt, Carello, & Turvey, 1990; R. C. Schmidt & O'Brien, 1997; R. C. Schmidt & Turvey, 1994; Tognoli et al., 2020).

Human behaviors unfold in ever-changing environments. These often include other individuals who produce actions in rhythmic and dynamic sequences which, in turn, provide affordances for interaction (Cummins, 2009a; Phillips-Silver, Aktipis, & Bryant, 2010; Richardson, Dale, & Shockley, 2008; Richardson, Marsh, & Baron, 2007). Central to this is the fundamental human skill known as sensorimotor synchronization (SMS), which underlies various forms of behavioral alignment with environmental rhythms. Broadly speaking, SMS refers to the temporal coordination of an action with a repetitive and, to a certain extent, predictable external event (Bruno H. Repp, 2005). It is realized via three functional processes: 1) the perception of timing in a rhythmic stimulus, 2) the production of rhythmic movements, and 3) the multisensory integration of perceived rhythms with produced motor rhythms (Phillips-Silver et al., 2010). In social contexts, the engagement of such functional processes by different individuals can lead to interpersonal synchrony, a collective state that emerges when two or more individuals mutually adjust towards a stable temporal relationship. This state is reached through a process of *dyadic entrainment* (Clayton, Jakubowski, & Eerola, 2020; Clayton, Sager, & Will, 2005; Knoblich, Butterfill, & Sebanz, 2011; Phillips-Silver et al., 2010; Phillips-Silver & Keller, 2012; Rosso, Maes, & Leman, 2021).

A view of joint-action informed by ecological psychology (Gibson, 2014; Marsh, Richardson, Baron, & Schmidt, 2006; Turvey, 1992) suggests a hierarchy of four mechanisms that facilitate social interaction. These range from the most basic, dyadic entrainment, to the more complex levels of intentional simulation, shared perception, and shared intentions (Knoblich & Sebanz, 2008). The hierarchy is structured such that basic interpersonal processes are embedded within more advanced functions that serve joint-action and communication, with entrainment as the most fundamental mechanism. Musical ensemble performance serves as a prime example of temporal coordination, such as when musicians' rhythms organize into complex patterns by engaging all these functions (P. E. Keller et al., 2014; Sebanz, Bekkering, & Knoblich, 2006). However, entrainment always operates below the level of intentionality, spontaneously drawing performers towards a synchronized state (Clayton et al., 2005; Marc Leman, 2016).

Empirical research has leveraged controlled experimental approaches to induce and study spontaneous forms of entrainment, including reduced tasks where participants are instructed to tap their finger while synchronizing with a rhythmic cue or with a human partner (for a review, see Bruno H. Repp & Su, 2013). Given such minimal forms of interaction, coordination dynamics has proven to be a valuable framework for understanding the underlying control mechanisms. These appear to be so general that dyadic entrainment manifests consistently across effectors and task contingencies (Issartel, Marin, & Cadopi, 2007; Miyata, Varlet, Miura, Kudo, & Keller, 2017, 2018; Oullier, de Guzman, Jantzen, Lagarde, & Kelso, 2008; Richardson, Marsh, Isenhower, Goodman, & Schmidt, 2007; Richardson, Marsh, & Schmidt, 2005; R. C. Schmidt et al., 1990; R. C. Schmidt & O'Brien, 1997). Traditionally investigated in the context of visuomotor coordination, the investigation of dyadic entrainment has subsequently expanded to interactions mediated by audition (Demos, Chaffin, Begosh, Daniels, & Marsh, 2012; Ole A. Heggli, Konvalinka, Kringelbach, & Vuust, 2019; Konvalinka, Vuust, Roepstorff, & Frith, 2010; Miyata, Varlet, Miura, Kudo, & Keller, 2021; Néda, Ravasz, Brechet, Vicsek, & Barabási, 2000; Nessler & Gilliland, 2009; Thomson, Murphy, & Lukeman, 2018; Xenides, Vlachos, & Simos, 2008) and mechanical coupling (Crombé, Denys, & Maes, 2022; Cuijpers, Den Hartigh, Zaal, & de Poel, 2019; Harrison & Richardson, 2009; Nessler & Gilliland, 2009; Zivotofsky & Hausdorff, 2007).

Informational coupling: a basis for social synergies

Temporal alignment in coordinated behavior necessitates the rhythmic units of a system to exchange information. Every interaction is realized on top of an informational basis,

which provides a structure for information flow and a substratum for feedback loops between the system components. The establishment of *informational coupling* among these units is a necessary condition and a fundamental prerequisite for the emergence of systemic behavior.

Intrapersonal coupling and bodily synergies

At the *intrapersonal* level, coupling is relatively straightforward, as it is inherently given by the human body and nervous system. Consider an individual coordinating the timing of two limbs for activities such as drumming, walking, or any synergetic movement. The ability of individual body segments to operate in concert relies not only on the brain, as the control center, but also on the physical coupling between body parts. This configuration allows information to flow through the physical contiguity of the human body via muscles, bones, tendons, and nerves. This bodily informational structure is largely consistent across various activities and over time, aside from major changes due to accidents. It is also continuous and persists irrespective of attention focus and fluctuations (R. C. Schmidt et al., 1990). This setup of physical linkage leads to a set of anatomical and functional constraints, which aid in solving the coordination problem posed to systems with numerous degrees of freedom (Latash, 2010). In other words, even the most ordinary of human activities is extremely complex in its structural organization. For this reason, the high dimensional state spaces of human behaviors must be compressed into lower dimensional systems in order to make them controllable.

The solution lies in the aggregation of individual variables into larger groupings known as *synergies*, which essentially are functional assemblies of structural elements (such as neurons, muscles, and joints), transiently bound to operate as a coordinated entity (J. A. Scott Kelso, 2009). During movement, the internal degrees of freedom are controlled indirectly, facilitated by anatomical and functional constraints which impose relatively fixed and autonomous relationships among the parts. This arrangement simplifies the system's control and sustains the integrity and stability of its behavior. At the same time, fluctuations are compensated by coordinated adjustments of its constituent parts in a function- or task-specific fashion (Latash, 2010). For instance, consider the act of walking: the control of individual muscles and the alternating pattern of the two legs can be delegated to the synergies of the locomotor apparatus, allowing the individual to consciously control only a few lower-dimensional parameters. Over-simplifying this concept, once the walking direction is set, gait control can be modeled as a single oscillator capable of adjusting its frequency to control the pace, potentially aligning itself with external oscillators like a musical beat (Bart Moens et al., 2014; Styns et al., 2007; van

Noorden & Moelants, 1999) or another walking individual (Zivotofsky & Hausdorff, 2007). The actual orchestration of all the muscles involved, and the postural adjustments in response to obstacles along the way, can be largely automatized and delegated to the synergy responsible for the locomotory function.

Interpersonal coupling and social synergies

The coordination problem becomes more complex in the case of *interpersonal* coupling, as individuals with independent nervous systems necessitate the exchange of information using one or more sensory channels (R. C. Schmidt et al., 1990). The two conditions satisfied for intrapersonal coordination of bodily synergies, namely the presence of a single control center (the brain) and a consistent informational structure (the body), do not exist at the interpersonal level. Thus, even minimal forms of dyadic behavior necessitate the emergence of implicit strategies (Ole A. Heggli et al., 2019; Ole Adrian Heggli, Cabral, Konvalinka, Vuust, & Kringelbach, 2019; Konvalinka et al., 2010), or more explicit joint planning when actions require a higher level of complexity (Knoblich & Sebanz, 2008). While anatomical connections facilitate coupling within body parts, coupling between bodies depends on a more diverse and intricate set of contingencies. For example, the medium, physical obstructions, relative positioning of bodies in physical space, and levels of attention can all impact access to mutual information. These factors, in turn, influence the information content resulting from the units, feeding the continuous action-perception loop underpinning the interaction (Phillips-Silver et al., 2010; Tognoli et al., 2020).

Dyadic entrainment can be seen as a dynamic mechanism through which information is shared between individuals (Hasson & Frith, 2016), where the spontaneous tendency of individuals in a group to mimic each other's actions promotes a basic spatiotemporal alignment among them. The coordinated duration, spacing, and phasing of movements is realized through a continuous mutual adaptation process (Konvalinka et al., 2010), leading to synchronized behavior among participants. Intriguingly, this alignment process shares commonalities with the mechanism responsible for the flocking behavior (Belz, Pyritz, & Boos, 2013; Biro, Sumpter, Meade, & Guilford, 2006; Couzin, 2009; GrÜnbaum, 1998) and the synchronous chorusing observed in a variety of animal species (Backwell, Jennions, Passmore, & Christy, 1998; Greenfield, 1994a, 1994b). The collective alignment of units within these groups, strictly reliant on an external signal, is governed by a few simple rules such as following the direction of one's nearest neighbor or minimizing the time lag between acoustic signals. Notably, the very notion of synergy expands to dyads or groups of human individuals who, as long as they are coupled, form a low-dimensional

reciprocally compensating system (Riley, Richardson, Shockley, & Ramenzoni, 2011). The emerging type of self-organized group behavior cannot be understood as a linear composition of its parts (Marsh et al., 2006), and can operate without the need for centralized control (Ballerini et al., 2008).

This brings us to introduce a core idea in the present work: *control structures emerge over informational structures* (R. C. Schmidt et al., 1990). In other words, how individuals coordinate with one another largely depends on how they exchange, access, and process mutual information. We posit that the physical properties of the information perceived by individuals (*action observables*) and the access to such information (*perceptual systems*) are two distinct dimensions of interpersonal coupling, which can be manipulated independently within an experimental design. This means that even if a perceptual system is optimized to access certain features of a stimulus (Comstock, Hove, & Balasubramaniam, 2018), it does not necessarily imply that these features are conveyed by the other person's actions, or that the physical contingencies of the environment allow that particular system to access and process them. This presents a picture of interpersonal coupling where multiple variables and conditions interact to create a dyadic system. This dissertation strives to cover some fundamental aspects of its complexity, through a series of empirical studies presented in the subsequent chapters.

Attractor dynamics

Synthesizing the evidence across scales and species, one notion becomes clear. While central control is not always necessary for organized behavior up to a certain level of complexity, the presence of informational coupling is indispensable. Information exchange is a necessary condition for any interaction. Determining whether it is also a sufficient condition constitutes a key research question across this dissertation.

The concept of *attractor dynamics* provides a foundational understanding of how rhythmic coordination functions at a systemic level. At the core of this understanding is the Haken-Kelso-Bunz (HKB) model, which frames rhythmic bimanual finger coordination as a system of coupled nonlinear oscillators, mutually attracted towards in-phase and anti-phase synchronization (Haken, Kelso, & Bunz, 1985). Over time, the HKB model has become an essential formal construct for explaining the establishment of stable patterns in intrapersonal rhythmic movements (Beek, Peper, & Daffertshofer, 2002). The seminal works on inter-limb coordination by Kelso and colleagues have expanded the scope of HKB to encompass rhythmic coordination within the same class of effectors, between different

effectors, and crucially, between two individuals (R. C. Schmidt & Turvey, 1994; Richard C. Schmidt & Richardson, 2008).

Central to this model is the concept of an *attractor*, a basic control structure underlying entrainment across rhythmic units (Marsh et al., 2009; R. C. Schmidt et al., 1990), which enables the system to self-organize spontaneously in a state where its trajectories converge and persist, maintaining stability (Tognoli & Kelso, 2014). In this stable state, the components settle on a common frequency, in a mode-locking defined by their phase relationship. Experimental procedures in coordination dynamics, some of which we expand upon in this work, aim to empirically define an *attractor landscape* (Schöner, Zanone, & Kelso, 1992; Tuller & Kelso, 1989, 1985; Yamanishi, Kawato, & Suzuki, 1980). This layout, operationalized by a collective variable expressing the temporal relationship between the components, presents a configuration of attractor points as preferred coordinative states of the system and quantifies their relative strength. Escaping one of these states or resisting its force in its proximity requires significant energy, an external perturbation, or intentional effortful control in humans. Nevertheless, the attractor landscape is not static: this can be reshaped and learned, offering a valuable framework for understanding how new motor skills can transition to automatic control through extended practice (Dhawale, Smith, & Ölveczky, 2017; Schöner et al., 1992).

Research by Schmidt and colleagues (R. C. Schmidt et al., 1990; R. C. Schmidt & Turvey, 1994) explored whether certain entrainment phenomena observed as within-person coordination also apply between persons, and whether the same general dynamical principles govern both instances. These studies, along with subsequent investigations, provided evidence that the basic control structure of an attractor underpins the frequency entrainment and phase locking between humans. As such, a dynamical interpretation of entrainment can be applied to the visually coordinated phasing of limbs between two individuals. The experiments confirmed the HKB prediction that a dyadic system would find stability in in-phase (aligned) and anti-phase (opposed) coordination patterns, with the former attractor being stronger than the latter. When the system is pushed to instability by increasing the movement's frequency beyond a certain threshold, the anti-phase attractor loses traction in favor of the in-phase pattern.

The existence of attractor dynamics governing dyadic entrainment implies that the principle of "control structures forming over informational structures" (R. C. Schmidt et al., 1990) can be translated into a specific working hypothesis: *in informationally coupled human dyads, the attractor dynamics are dependent on how information is accessed and processed within the dyad*. The systematic investigation of this idea will necessitate a method to quantify the strength of attractors and a strategy to manipulate informational

coupling between individuals. Both aspects will be addressed in the section 'Major methodological concepts' and will form a significant part of this dissertation.

Major theories

Coordination dynamics offers an elegant framework for understanding the execution of complex behavior in biological systems, including humans. However, it can be argued that its explanatory power is mainly confined to the descriptive level. In contrast, cognitive and social neurosciences aim to move beyond this level and provide explanatory accounts for human rhythmic interactions. Since there is no consensus yet on the nature of the underlying mechanism, we hereby introduce two major contemporary debates in the field, which will be revisited in the discussion of our empirical findings.

Explaining synchronization: predictive coding vs dynamical systems

The first debate prescinds the social aspect of interactions. It focuses instead on the principles of anticipatory mechanisms coming into play when humans synchronize with environmental rhythms. Anticipatory synchrony, a well-documented phenomenon in human SMS, is necessitated by the inherent delay in movement preparation (G. Aschersleben & Prinz, 1995, 1997; Gisa Aschersleben, 2002; Dunlap, 1910; Johnson, 1899; Bruno H. Repp, 2000). This delay compels the anticipation of future perceptual events, enabling a timely preparation and execution of movement to achieve synchronization (Palmer & Demos, 2022). Notably, even sensorimotor experts such as musicians exhibit the same phenomenon, albeit with a marginally improved accuracy (Bruno H. Repp & Su, 2013). Furthermore, effective SMS requires a certain degree of predictability in the temporal structure of the stimulus, which in turn enables successful anticipatory alignment of movement with external rhythms.

Predictive coding and *dynamical systems* theories emerge as the primary theoretical contenders in explaining anticipatory synchronization, and both of these frameworks adhere to the free energy minimization principle (Bruineberg, Kiverstein, & Rietveld, 2018). This principle denotes the drive of adaptive living systems to minimize free-energy in their interactions with the environment (K. J. Friston & Stephan, 2007), essentially resisting the tendency towards disorder and entropy of sensory states (K. Friston, Kilner, & Harrison, 2006). This minimization process can be achieved in two ways: either by predicting or anticipating incoming sensory input, or by modifying the environment to

align with expected outcomes. To anticipate effectively, an organism must be well-adapted to its ecological niche. This ensures that the interrelated dynamics between the organism and its environment stay within a limited range of states that support the organism's survival in its specific habitat (K. Friston, 2011).

Although both theories share this principle as a common ground, they diverge in explaining its neural implementation. The Bayesian formulation of predictive coding postulates that neural activity is inherently predictive (K. J. Friston & Stephan, 2007; K. Friston & Kiebel, 2009). In this view, the brain is constantly generating top-down predictions about the environment, updating its model based on new bottom-up sensory evidence, and minimizing eventual mismatches (Clark, 2017; Hohwy, 2013). As per this perspective, free-energy essentially corresponds to the amount of prediction error (K. Friston & Kiebel, 2009).

Contrarily, dynamical systems theory leverages self-organization principles within agent-environment interactions, modelling synchronization behavior in terms of non-linearly coupled oscillators described by differential equations (Assisi, Jirsa, & Kelso, 2005; Fink, Foo, Jirsa, & Kelso, 2000; Schöner & Kelso, 1988; Torre & Balasubramaniam, 2009). This approach can notably explain anticipatory behavior in basic rhythmic interactions without referencing internal models (Demos, Layeghi, Wanderley, & Palmer, 2019; Roman, Washburn, Large, Chafe, & Fujioka, 2019; Stepp & Turvey, 2010). However, it often lacks clear definitions of its neural underpinnings (Palmer & Demos, 2022). Broadly speaking, dynamic systems theory posits that intrinsic neural oscillations entrain to periodicities in rhythmic environmental signals (Large & Jones, 1999; Large & Snyder, 2009), and holds that the physical properties intrinsic to the neural system account for anticipatory and more general forms of synchrony (Large & Snyder, 2009). Within this framework, the coupling between internal oscillations and environmental rhythms explain the accuracy and consistency of SMS (M. C. M. van der Steen & Keller, 2013). Simply put, the model defines the strength of the entrainment by a specific coupling term.

To elucidate our stance in this debate, let us distill our perspective into a few key points:

- We value the predictive coding theory for its ability to address the role of the brain in generating top-down predictions about the environment, adjusting its model based on new bottom-up sensory evidence.

- Yet, we find the dynamical systems theory's approach compelling, for its ability to explain anticipatory behavior in basic rhythmic interactions without referencing internal models.

- Overall, our investigation leans towards the dynamical systems theory, due to its parsimonious explanations and the emphasis it places on agent-environment interactions rather than on brain-centric processes.

In and out of the brain: where is the foundation of social cognition?

A parallel debate in social neuroscience revolves around the fundamental question of how humans understand other minds during social interactions. Interactionists from the enactivist tradition assert that the foundation of social cognition rests in their coupling (Schönherr & Westra, 2019). To understand human social interactions, they argue, one must look beyond individual brains and their shared representations of the world, instead exploring the intrinsic properties of a higher-order dynamical system that emerges from the causal interdependencies of autonomous agents (De Jaegher, Di Paolo, & Gallagher, 2010). This principle traces back to Gallagher's seminal critique of traditional social cognition theories such as the theory of mind and the simulation theory, which rely on cognitive abilities to explain and predict others' mental states and behaviors (Gallagher, 2001). These perspectives typically employ observation paradigms that emphasize individual minds in social interactions. A growing consensus suggests that interaction theory and its corresponding methods, which approach social cognition at the collective level of group interactions, should be further developed to complement existing theories and foster a more comprehensive understanding of social cognition (Schönherr & Westra, 2019).

In this debate, our position can be summarized as follows:

- We value the arguments of the enactivist tradition, asserting the foundation of social cognition in coupling and stressing the importance of exploring the intrinsic properties of higher-order dynamical systems.

- However, we believe it is crucial to not disregard the role of the brain in this process. Even with the interactionist approach, there is an undeniable element of perceptual input processing happening at the central level.

- Our stance is thus an attempt to merge these perspectives, acknowledging the need to examine agent-to-agent and agent-to-environment couplings, but also recognizing the mediating role of the brain in these interactions.

Our theoretical stance: a dynamic balance

Where does this dissertation stand with regard to these debates? While there are sections in the following chapters in which we lean towards a specific theoretical stance, we stress here the importance of a balanced perspective when interpreting our empirical findings. In general, we value the deflationary and parsimonious nature of dynamical systems, and favor an enactive systemic approach when it comes to explaining dyadic interactions (Bruineberg et al., 2018; Schönherr & Westra, 2019). We reiterate that we employ a very low-level, general definition of interaction. In the empirical studies presented throughout the dissertation, humans interact with each other deploying all the specifics of the human body and nervous system. Yet, we consider interactions at a low level that prescinds intentionality, viewing humans as general rhythmic units in a coupled system. Other authors may use a more stringent definition, arguing that an intentional stance is necessary when discussing the concept (Schönherr & Westra, 2019). While we believe that our research has implications for higher-level social cognition, our main concern lies in explaining the fundamentals of interpersonal synchronization.

We posit that shifting the focus outside of the brain, towards agent-to-agent and agent-to-environment couplings, is a convenient approach for modelling human interactions. We propose that both dispositional factors (personality traits, sensorimotor abilities, attitudes, and intentions towards the partner and the interaction) and situational factors (similarity and liking for the partner, evaluation of the environment) can be modeled as modulators of the coupling between interacting units in a brain-body-environment system. However, we acknowledge that moving between 'in' and 'out of the brain' is a matter of levels of analysis. We cannot deny the role of the brain as a central controller at the individual organism scale, and hence its mediation of the interaction between human agents. Even the most radical embodied perspective would admit that some form of processing of perceptual inputs occurs at the central level. This is undeniable and is rooted in the neuroanatomy of sensorimotor pathways. What happens between perceptual input and motor output, is where major theories differ.

More than oscillators

We humans are clearly more than oscillators. Our nervous systems are capable of complex information processing, they hold internal representations of the world and can execute elaborate intentional planning. However, it is fascinating to observe how the emergent properties in minimal rhythmic behavior resemble those in physical systems. For instance, the replication of the classic Huygens' experiment (Huygens, 1888) with humans using

mechanical coupling (Crombé et al., 2022) affirms the framework's validity in describing how coupled rhythmic behavior results in spontaneous synchronization between individuals, and in explaining its basic mechanisms (Beek et al., 2002; de Poel, 2016; Haken et al., 1985; Lagarde, 2013).

We assert that such a framework does not conflict with minimal forms of self-other internal representations. These have been effectively modeled as coupled action-perception units within and between brains (Ole Adrian Heggli et al., 2019; Ole Adrian Heggli, Konvalinka, Kringelbach, & Vuust, 2021), and are mirrored in the topographical arrangement and functional connectivity of neural subsystems dedicated to representing the 'self' and the 'other' (Ole Adrian Heggli, Konvalinka, Cabral, et al., 2021). Let us consider the paradigmatic example of two individuals facing each other, while moving their index finger up and down (Tognoli et al., 2020). The spatial displacement of one finger over time is first registered onto the partner's retina. The resulting oscillatory retinotopic mapping propagates from peripheral receptors to central sensory areas and eventually feeds into the perceiver’s motor system, which in turn sends efferent signals to move their finger. We argue that this internal mapping constitutes a basic form of representation, with a clear segregation between information flows generated by the 'self' and the 'other'. The interaction between the streams results in motor interference solved by aligning in both temporal and spatial dimensions (J. M. Kilner, Paulignan, & Blakemore, 2003). Besides their kinematic profile, the geometry and identity of bodily effectors are somatotopically represented in sensorimotor cortices. These representations constitute the fundamental elements for plastic representations of the bodily self (Tsakiris, 2010, 2017), which can be dynamically updated based on situational contingencies and facilitate processing of self-other discrimination from the early stages of perceptual processing (Galigani et al., 2021).

The point we aim to convey is that rejecting the notion of internal models to explain basic adaptive interactions with the environment and other human agents doesn't entail entirely rejecting representations. Our position may be defined representationalist in a broad sense. The basic observation of evoked-responses as changes in neuronal excitation in response to an event (M. X. Cohen, 2014; Luck, 2014), indicates that the brain processes sensory information, encodes it into neural patterns, and uses these patterns to form a neural representation of the environmental events. We argue that even a stream of passive bottom-up responses evoked by isochronous stimulation is, by definition, a neural representation of environmental rhythm, and a marker of the processing of low-level features of the stimulus (Norcia, Appelbaum, Ales, Cottereau, & Rossion, 2015; Novembre & Iannetti, 2018; Vialatte, Maurice, Dauwels, & Cichocki, 2010). Taking music as a paradigmatic case of a temporally structured perceptual stimulus, and considering rhythm

as a crucial dimension of it, it is not surprising that profound insights into representation and processing at the central level are found in music neuroscience (Vuust et al., 2022). The close connection between rhythmic structure in audio signals and neural signals has been highlighted from both fundamental and signal processing perspectives (Large & Snyder, 2009; Lenc et al., 2021; Nozaradan, Peretz, Missal, & Mouraux, 2011). Specifically, prominent periodicities and hierarchical structures can be mapped onto brain activity as recorded by magneto-electroencephalography (M/EEG), enhancing our understanding of how the brain centrally encodes and represents rhythmic structures.

As discussed above, environmental and human rhythms are not passively represented. The internal organization of brain physiology is predisposed to adaptation, when coupled to changing environments through perception. Adaptive interactions necessitate underlying neural mechanisms capable of tracking changes and enabling flexible behavior (Criscuolo et al., 2022), guided by predictions which are thought to be encoded in the temporal scaffolding of frequency-specific oscillatory dynamics (Nobre & van Ede, 2018). This adaptation process is the reason why neural entrainment, namely the unidirectional synchronization of neural oscillations to an external rhythmic stimulus (Haegens & Zion Golumbic, 2018; Lakatos et al., 2019), is a process of substantial significance for the present work. We believe that explicitly modelling the interaction mechanisms between internal and external rhythms is key to the understanding of coupled behavior. From an embodied cognition perspective, the intrinsic physical properties of the nervous system cannot be neglected. In certain instances, their interaction with the environment can result in emergent properties which alone can explain behavioral phenomena (Large & Snyder, 2009). In other instances, it can provide a neural substratum to putative theoretical constructs (Heilbron & Chait, 2018). Either way, this level of analysis is an integral part of this dissertation, as the link between the theories and the underlying physiological processes. To summarize:

- We adopt a representationalist position in a broad sense, acknowledging that our brains are not passive receivers of information but actively encode sensory inputs to form neural representations of environmental events.

- We believe that representations do not necessarily require the assumption of complex internal models. Simple neural representations of environmental rhythms can provide a substratum for interaction mechanisms with intrinsic dynamics (e.g., via neural entrainment).

- The explicit modeling of interaction mechanisms between neural and environmental rhythms is key to understand coupled behavior.

The picture of the brain emerging from our discussion is that of an intrinsically social instrument, whose overall functioning exhibits the property of adaptation to the environment. This adaptive property underscores the social nature of humans and suggests that basic sensorimotor processes, such as SMS, form a predisposition towards sociality. Our investigation into human rhythmic interactions operates within the context of this adaptation process. Its principles are to be found at the level of behavioral and neural collective dynamics and the conditions which enable and support these dynamics.

In conclusion, this introduction has outlined the major theoretical concepts that will be explored in this dissertation. However, it is important to note that proving or disproving overarching theories of human brain and behavior falls beyond the scope of this work, given that the methodologies of our studies were not designed for such a purpose. Consequently, we abstain from favoring any specific theory in our interpretations unless the conclusions are firmly grounded in our data.

Major methodological concepts

Eliciting and quantifying dynamics

The overarching rationale of the methodologies developed and used throughout this work is to elicit, quantify, and model the dynamics of human rhythmic interactions. Therefore, our primary focus is on the temporal structure and the changing nature of signals generated by rhythmic events of interest: stimuli, behavior, and neural activity. Not only must our methods be sensitive to these temporal features, but they must also capture the interactions across signals as they evolve over time. In the following chapters, these methods will be applied to dyadic and individual experiments. The former involves genuine interactions in the physical presence of a human partner, which is the primary subject of this work. The latter involves synchronization tasks performed in isolation, with dynamic computer-generated rhythmic cues.

While the unpredictability of social interactions provides a valuable resource for the emergence of the phenomena in our interest, it carries the obvious limitation that we have limited experimental control over the unfolding of the interaction and over the stimuli when they are human-generated. In a dyadic interactive context, part of the stimulation for the participants consists of the partner's actions, which are largely variable, unpredictable, and beyond the experimenter’s control. Therefore, we believe that our

research needed to be supplemented with studies involving individual participants, undergoing more controlled procedures to help us gain further insight into the behavioral and neural entrainment to controlled dynamic stimuli. This complementary approach was designed to address general sensorimotor synchronization principles that prescind the social dimension. Regardless of this distinction, each of the experiments presented in this dissertation was informed by the following guiding principles:

- *Rhythmicity*. The aim of stimulation in the context of a task is to induce rhythmic activation in the participant(s), measurable at both behavioral and neural levels.

- *Interactivity*. Moving beyond traditional stimulus-response paradigms prevalent in experimental psychology, participants are engaged in a scenario where their bodily rhythmic activation continuously interacts with dynamic rhythmic stimulation.

- *Time-variance*. The interaction between rhythms generates an emergent temporal structure, which changes as it unfolds over time. The dynamics underlying these changes are our primary object of analysis.

- *Systematicity*. While our experiments are designed to allow for dynamics and variability within tasks, they also aim to uncover general processes underlying rhythmic interactions. Therefore, they were designed with the goal of enabling organized, methodical, and consistent measurements. Two key notions in this regard are:

    - *Ground truth*. This refers to the temporal structure of the stimuli, providing a clear picture of the system's expected behavior under ideal conditions and in the absence of noise. Numerically, this offers a reference point for assessing systematic variations in the system due to experimental manipulations.

    - *Control parameter*. This refers to an independent variable that is systematically manipulated to drive the behavior of the system under analysis. This will be presented in greater detail in the upcoming section (see 'Dyadic studies').

Given the need to induce the dynamics of interest in the behavioral and neurophysiological activation of the participants, it is necessary to have at one's disposal a set of signal processing tools to quantify them. Due to the cyclical nature of the processes

under investigation, these can be approximated with reasonable accuracy within a framework of coupled oscillators. We define an oscillatory signal as a periodic, recurring time series defined by its amplitude, frequency, and phase. It can be modeled as a sinusoidal process using the following mathematical equation:

$$y = A * sin(2\pi * f * t + \varphi)$$

where:

- *Amplitude* (A) quantifies the displacement of the waveform from its equilibrium point. It corresponds to the height of the wave's peaks or the depth of its throughs, and determines the deviation of the oscillating system from its rest position.

- *Frequency* (f) quantifies the number of cycles or oscillations that a sinusoidal waveform completes in one second. It is expressed in hertz (Hz), where 1 Hz corresponds to one cycle per second. It determines the repetition rate of the sinusoidal process, with higher frequencies resulting in more cycles per second. It is the reciprocal of the period, which is the time taken to complete one full cycle.

- *Phase offset* (φ) quantifies the initial position or angle of the sinusoidal waveform at time t=0, typically measured in degrees (°) or radians (rad). The phase determines the horizontal shift of the sinusoid relative to a reference point, such as the origin. A change in the phase can cause the waveform to start at a different point in its cycle, effectively changing the position of the peaks and troughs along the time axis.

Given these basic elements for modelling the activation of rhythmic signals, we require a toolkit of analytical methods to quantify their interactions. Cross-frequency coupling is a term encompassing various co-variational phenomena in the dynamics of multiple signals oscillating at different frequencies, or within different oscillatory components of a broadband signal. Among the different instances of this interaction, we focus on amplitude modulation (AM) and frequency modulation (FM), where a low-frequency signal acts as the modulator and a high-frequency acts as the carrier signal undergoing modulation. These concepts will form the methodological backbone of Chapters 5, 6, and 7, which delve into the neural mechanisms underlying behavioral entrainment.

Entrainment is not explicitly classified as a type of cross-frequency coupling. Although both phenomena involve interactions between oscillatory signals, and both can be observed in neural processes (Allen et al., 2011; Canolty et al., 2006; Lakatos et al., 2019), cross-frequency coupling encompasses a broader range of interactions. Entrainment, on the other hand, specifically refers to a type of interaction where the phase of one oscillating signal aligns with another which acts as the driving rhythmic force, until their phases and frequencies lock over time (Lakatos et al., 2019). A state of phase-locking implies a relative-phase of 0 rad between two signals, whereas frequency-locking implies consistently maintaining the same frequencies and a fixed relative phase. When the two signals interacting via entrainment oscillate at different frequencies, the initial gap decreases until synchronization is achieved. This phenomenon was elegantly modeled by the Kuramoto in a set of differential equations which describe the collective behavior of a large number of interacting oscillators (S. H. Strogatz, 2000):

$$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j=1}^{N} \sin\left(\theta_i - \theta_j\right), \qquad i = 1, \dots, N$$

where:

- *Angular position* (θ_i) quantifies the phase of the *i*th oscillator. It is typically measured in degrees (°) or radians (rad) and represents the state of the oscillator in its cycle. $\dot{\theta}_i$ is the first derivative of the phase and refers to its velocity or rate of change.

- *Natural frequency* (ω_i) quantifies the intrinsic tendency of the *i*th oscillator to oscillate at a certain rate.

- *Coupling strength* (*K*) quantifies the degree of interaction between the oscillators in the system. It determines the extent to which the oscillators influence each other's behavior.

- *Number of oscillators* (*N*) represents the total count of oscillators in the system, as the model scales to an infinite number of interacting oscillators. The term serves to adjust the coupling strength for the total count.

- Summation ($\sum$) represents the sum of the sine of the phase differences between the *i*th oscillator and all other oscillators *j* in the system. This term captures the

influence of the other oscillators on the i^{th} oscillator and is responsible for driving the synchronization process.

In essence, the phase velocity of each oscillator is a sinusoidal function of the phase difference with all other oscillators in the system, scaled by the coupling term and adjusted for the total number of oscillators. Since $\sin(\theta_i - \theta_j) = 0$ when $\theta_i - \theta_j = 0$ or when $\theta_i - \theta_j = \pi$, the in-phase and anti-phase states represent stable points where oscillators settle and maintain a common frequency.

While the Kuramoto model will not be explicitly used to analyze data from our studies, it provides valuable insight. Specifically, it highlights the relationship between phase and frequency of the oscillators and implies that there are attractor dynamics driving towards phase- and frequency-locking. These are all core methodological concepts in this dissertation, which will be recurrent throughout the chapters. The model identifies in-phase and anti-phase patterns as the states where oscillators settle their phase at a stable point and, consequently, synchronize their frequencies. Importantly, entrainment involves adjustments of the signal's phase, meaning that the oscillation would shift its phase to speed up or slow down until synchronization is achieved. It is noteworthy that altering the velocity of the phase is equivalent to changing its instantaneous frequency.

These concepts will be discussed extensively and in greater detail in the studies dedicated to neural dynamics (Chapters 5-7). As a final note, when discussing neural entrainment, we should recognize that we are making specific assumptions about brain function. Primarily, we are assuming that brain oscillations are 'real'. That is, neural activity is characterized by ongoing, endogenous oscillatory activity that does not require to be triggered by external stimulation. Additionally, these oscillations should have the flexibility to adjust their frequency over time to synchronize with the stimulation. Therefore, the search space of our analysis is guided by these assumptions and is constrained within an oscillatory framework.

Dyadic studies

We now move on to the methodology encompassing a series of studies dedicated to dyadic interaction (Chapters 2-4). While most of these aspects will be revisited in the 'Methods' sections of the individual studies, we take this opportunity to familiarize the reader with the rationale and the principles guiding the design of a cohesive series of experiments, and to highlight some nuances and implications that are not included in the

published manuscripts. What follows is the presentation of the building blocks of the dyadic studies: experimental *setup*, *task*, *paradigm*, and *design*. Each block is necessary for the next one in order to carry out the experiments.

Experimental setup

This block encompasses the set of physical conditions and constraints that enabled us to conduct the experiments. Here are the essential elements needed to investigate dyadic rhythmic interactions:

a) *Dual acquisition system*. At the very least, two devices are needed to record the timing of behavioral responses from each partner in the dyad. Pressure sensors placed inside two pads and connected to the same voltage microcontroller represent a practical and portable solution, which we extensively adopted to detect the onsets of hitting time with sub-millisecond resolution. Motion capture offers a more sophisticated acquisition method to sample, at a fixed rate, the position of infrared reflective markers placed on body parts of interest. This technique offers advantages such as introducing the 3D spatial dimension in the analysis and the application of timeseries analysis without the need for interpolation. However, it also has drawbacks such as the dependency on a more sophisticated infrastructure, a more cumbersome setup, the need for additional data pre-processing, and potential incompatibility of sampling rates with other acquisition devices. In addition to behavior, we are interested in recording physiological responses from the participants while they are engaged in the experiment. Regardless of the specific parameter under investigation, it is crucial that the sampling rate of the recording device is high enough to capture the frequency of the dynamic phenomenon of interest. Electrophysiological tools can typically record fast-changing signals such as electrocardiogram (EKG), electrooculogram (EOG), or electroencephalography (EEG). In our setup, we integrated dual-EEG recordings to record brain activities from participants and quantify the neural dynamics underlying interpersonal coordination. The simultaneous acquisition of neuroimaging techniques is often referred to as 'hyperscanning' in the literature (Zamm et al., 2023), and we will use this term throughout the text.

b) *Dual stimulation system*. Unless the experiment explicitly requires participants to improvise an interaction or spontaneously produce and self-pace behavior, it is necessary to pace their behavior with rhythmic cues. It is preferable to provide

independent channels for the participants, which allows the flexibility to differentially manipulate the stimuli within the dyad. In practice, earplugs or earphones are preferred over speakers for auditory cues, while detached lights or separate headsets are preferred over a common monitor or projector for visual cues. While delivering stimuli to the participants, the stimulation system must keep logs of the data that generated the stimuli, particularly the timing of their presentation. These data will be used as the references for behavioral and neural timeseries.

c) *Synchronization system*. Regardless of the specific method, all analyses deal with temporal relationships between incoming data. To account for these cross-relationships among data streams, all devices need to be synchronized and maintain temporal alignment for the entire duration of the experiment. Six et al. (*in preparation*) are currently working on an extended technical discussion of all devices integrated into the ASIL infrastructure where all studies were carried out (Van Kets et al., 2021), the specific technical challenges of dealing with different sampling rates, and a general framework for tackling them. For the studies presented in present work, a simpler approach was tailored to the integration of the dual-EEG system in the facility, offering the advantage of portability across laboratories. The general synchronization schema is shown in *Figure 1.1*, with case-specific details presented in the 'Methods' section of each study.

d) *Coupling manipulation*: Finally, having discussed how perceptual coupling between participants is foundational to the present work, we need a way to experimentally control and manipulate it. The basic manipulation, common to most studies, is to set the coupling to 1 (enabled) or 0 (disabled) levels. Regardless of the contingencies, we want the ability to enable and disable the exchange of information between the participants, so that we have at our disposal control conditions where individual behaviors are uncoupled and therefore not mutually affected. Uncoupled conditions are crucial in our experimental designs because they provide a baseline to assess the effect of enabling the coupling. Alongside this basic 'ON/OFF' switch, the manipulation of specific dimensions of informational coupling will complete the *experimental design* of the presented studies.

Experimental task

Joint finger-tapping was the main task for our investigation. This has been a widely used method in interpersonal coordination studies (Bruno H. Repp & Su, 2013) due to

fundamental and practical reasons. The fundamental advantage lies in its simplicity: it offers a minimal form of rhythmic behavior that enables the investigation of temporal coordination between individuals, while maintaining a high degree of experimental control and isolating rhythm from other dimensions. From a practical perspective, joint finger-tapping, compared to full-body movements, is relatively easy to perform, less reliant on individual differences in body size, shape, and mobility, and allows for larger amounts of data collection due to less fatigue and higher production rate (Kliger Amrani & Zion Golumbic, 2022; London, 2012). Notably, minimizing body movements enables the collection of reliable electrophysiological data with techniques that are extremely sensitive to motion artifacts, such as M/EEG (Vidal, Rosso, & Aguilera, 2021). While this choice comes at the expenses of some ecological validity, the benefits outweigh the drawbacks when feasibility and reliability are prioritized. This holds particularly well when the setting allows to isolate fundamental properties transversal to different classes of natural behavior.

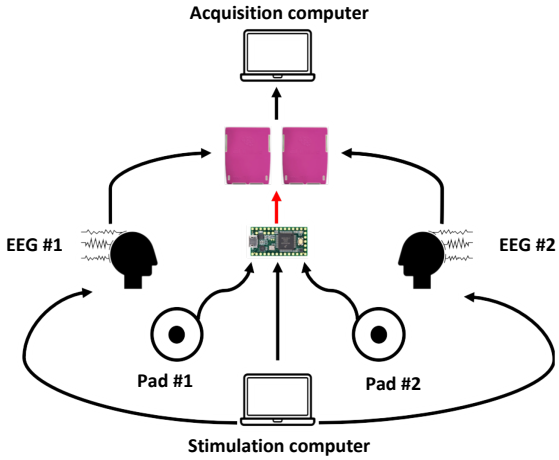


Figure 1. 1. Synchronization of devices in the setup. The Teensy 3.2 microcontroller, illustrated at the center of the diagram, serves as the primary serial/MIDI hub in the setup. From bottom to top, we have the following components. *Stimulation computer*: Ableton® Live 10 or Max/Msp were used as the main interface for presenting auditory stimuli to the participants. As the stimuli were generated based on MIDI signals, these could be dispatched to the Teensy for logging in a data stream synchronized with the pair of sensors used for collecting behavioral data. *Microcontroller*: Based on events of interest, the microcontroller dispatched TTL triggers via a BNC connector to mark events in the dual-EEG recording, facilitating offline alignment with the dual-stimulation and dual-acquisition systems. Triggering is represented with a red arrow. *Dual-EEG Systems*: Two eego™mylab systems were cascaded to simultaneously record 2x64 channels from two interacting subjects on the same *acquisition computer*. The same schema applies when scaling down the setup to a single participant. This scheme was previously presented at the ANT Neuro Meeting 2020 (Beune, FR) and the Sysmus 22 workshop (Ghent, BE).

Experimental paradigm

A task itself isn't particularly useful unless embedded within an experimental paradigm, which we define as the set of conditions allowing the researcher to isolate and investigate a particular phenomenon of interest. In our case, we aim to set the conditions for the emergence of coordination dynamics between partners, necessitating real interactions in physical presence for a relatively extended period of time. Such a scenario poses methodological challenges as the participants' movements act as dynamic stimuli for each other, influencing behavior and feeding a dyadic action-perception loop during the task (Phillips-Silver et al., 2010). Despite the limited control over the stimuli and the limited predictability of the responses, the interaction in its complexity is precisely the phenomenon of interest, which is the reason why many researchers have called for a 2nd person approach to the neuroscience of human behavior (Gallagher, 2001; Schönherr & Westra, 2019). We faced a difficult tradeoff: how and to what extent could we set constraints for an interaction to unfold freely while still being capable of measuring its dynamics?

Our solution was guided by dynamical system theory principles and inspired by early literature on coordination dynamics (J. A. S. Kelso, 1995; Tuller & Kelso, 1989). First, we identified a *collective variable* to quantify temporal relationships between partners, describe the dyad's collective behavior, and return a scalar value quantifying the degree of coordination within the system. This is known as the *order parameter* (Φ) (Tognoli et al., 2020). For instance, the relative phase between two rhythmic movements is a simple measure traditionally used to describe coordinated behavior. Throughout our studies, we will adopt a more complex yet related measure serving the same purpose. Second, as the dynamics of the interaction are our object of study, a global quantification of coordination was insufficient. We needed our collective variable to be time-variant, as information on dynamics is contained in its changes over time.

In order to elicit predictable patterns of variance over time, we required a method to drive the dyadic system through a space of coordinative states in a systematic and controlled fashion. This required the implementation of what is known as a *control parameter* (Ψ) (S. Strogatz, Friedman, Mallinckrodt, & McKay, 1994), a continuously manipulated independent variable. Independent variables are often conceived of in categorical terms as multi-level factors. To elucidate the difference, consider the following metaphor: imagine a control room from where we run our experiments, manipulating our independent variables through buttons and switches. Some can be toggled between 0-1, acting as a binary switch to enable or disable the variable, while others allow a selection from multiple levels: 1, 2, 3, and so on. These factors operate as switches in this metaphor.

The control parameter, conversely, functions more like a dial that the experimenter can continuously adjust, fine-tuning the value on a continuous scale during the experiment. This allows the investigation of collective behavior as it shifts in response to changes in the control parameter over time.

This dynamic approach forms the core of what we termed the ***drifting metronomes paradigm***. During a joint-finger tapping task, each member of the dyad is assigned a metronome and instructed to synchronize with it. A minimal gap in the metronomes' frequencies results in a constant accumulation of phase lag. By setting the metronomes to start in-phase, their collective behavior results in a linear de-phasing pattern. Step by step, their relative phase increases from 0 to π and subsequently decreases from π to 0 before the cycle repeats. The pattern is illustrated in *Figure 1.2*, and an audio sample of the pattern is provided in the 'Supplementary materials' of Chapter 2.

While partners cannot perceive the other's metronome, they are exposed to each other's movements under different conditions. Importantly, they are instructed to ignore the other person while maintaining synchronization with the assigned metronome. This results in a temporal mismatch between the assigned metronome and the partner's movement, which varies systematically as the metronomes de-phase. The paradigm's rationale is that, in absence of mutual influence between the partners, the collective behavior would track the deterministic trajectory of the metronomes, plus some error due to human motion variability. On the other hand, any systematic deviation from this pattern can only be explained by mutual influence due to the informational coupling.

The relative phase between the metronomes represents the control parameter of the drifting metronomes paradigm. It is mapped 1:1 onto time because we know the expected behavior of the system at every moment. In essence, this is what we previously described as ground-truth in the context of this paradigm. The emergence of consistent deviations from the ground-truth points at the influence of an *attractor*, whose position and strength are quantified via our procedure. The exploration of the whole space is expected to return the attractor landscape of every dyad (see *Figure 1.3*).

We list here some fundamental implications of the paradigm that will recur throughout all studies:

1) When dyadic entrainment occurs, it is *spontaneous* and against instructions and intentional attempts to ignore the partner.

2) When dyadic entrainment occurs, it is always at the expense of synchronization with the metronome. In other terms, the *coupling strength is distributed* across the metronome and the partner.

3) Dyadic entrainment is characterized by a dynamic *cooperation-competition* balance, or metastable behavior (Ole Adrian Heggli, Konvalinka, Kringelbach, et al., 2021; Tognoli & Kelso, 2014) where transitions from one state to another are promoted by attractors with opposite tendencies (Marsh et al., 2009). Note that the task is competitive in nature, as each individual actively tries to maintain the assigned rhythm.

4) Recurrent temporal structures in collective behavior as a function of the control parameter reveal the *attractor landscape* of the dyad (Schöner et al., 1992; Tuller & Kelso, 1989, 1985; Yamanishi et al., 1980).

These points hold provided the coupling between individuals is enabled, which brings us to the final building block of our methodology.

Experimental design

This is the stage dedicated to hypothesis testing. Under the overarching working hypothesis that the dynamics of attractors in coupled human dyads depend on how information is accessed and processed, each experiment is designed to manipulate different dimensions of informational coupling. With this foundation common to all dyadic studies presented in this dissertation, we now delve into the specifics of the experimental design.This is the stage where each study stands unique, setting the conditions to manipulate the independent variables of interest and assess their impact on the collective dependent variable. It is important to note that our dependent variable is time-variant, which must be considered and explicitly modeled in data analysis.

Returning to our control room metaphor, having explained how we control the 'dial' of drifting metronomes, we now turn our attention to the 'switches'. The first switch controls the coupling, toggling it between ON/OFF states. To statistically affirm that deviations from the drifting metronomes are explained by the coupling, above the chance level of random fluctuations, it is crucial to have a control condition where participants undergo the same procedure without perceiving each other. This provides a baseline for dyadic behavior in the absence of coupling, which is expected to follow the linearly de-phasing trajectory while accounting for random errors in individual synchronization performances.

Statistically contrasting a coupled condition against this baseline is more elegant and informative than running permutation tests or contrasting against the ground-truth of the metronomes' trajectory.

In tandem with the ON/OFF manipulation, we operationalize a different fundamental aspect of informational coupling in every study, manipulating it across two levels. As a result, all studies presented in this work adopt 2 x 2 within-subjects factorial designs, wherein all participants belong to the same group (young, healthy, non-professional musicians) and experience the full range of experimental conditions resulting from the combination of factors' levels. The decision for balanced designs facilitated the use of parsimonious mixed-models, eliminated the need for post-hoc comparisons, and ensured unambiguous interpretation of results.

We acknowledge that the variables investigated here are far from being exhaustive, and many more aspects of relevance are yet to be explored. Potential future directions will be addressed in the final chapter of this dissertation.

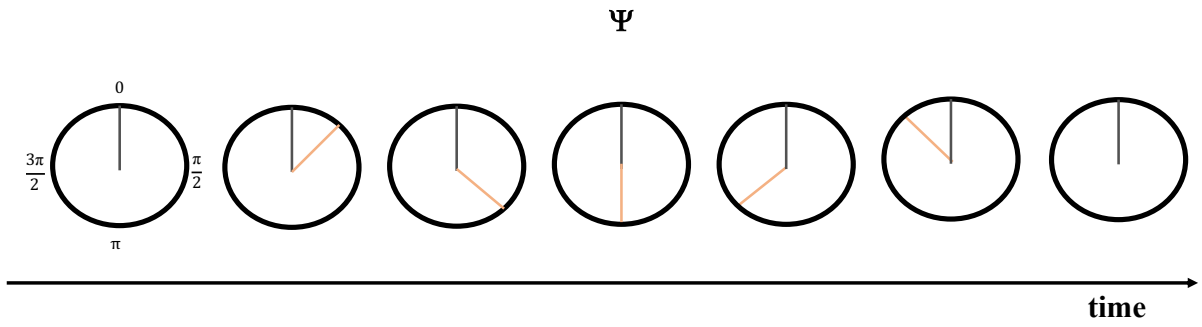


Figure 1. 2. The drifting metronomes cycle. Conceptual representation of the two metronomes' de-phasing pattern, over a full cycle. Note that the relative phase of the metronomes (Ψ) is mapped onto time, such that we know the expected state of the system at each point. In the actual implementation, it takes the metronomes 64 steps to complete the cycle in 39 seconds. The cycle repeats through 10 consecutive iterations.

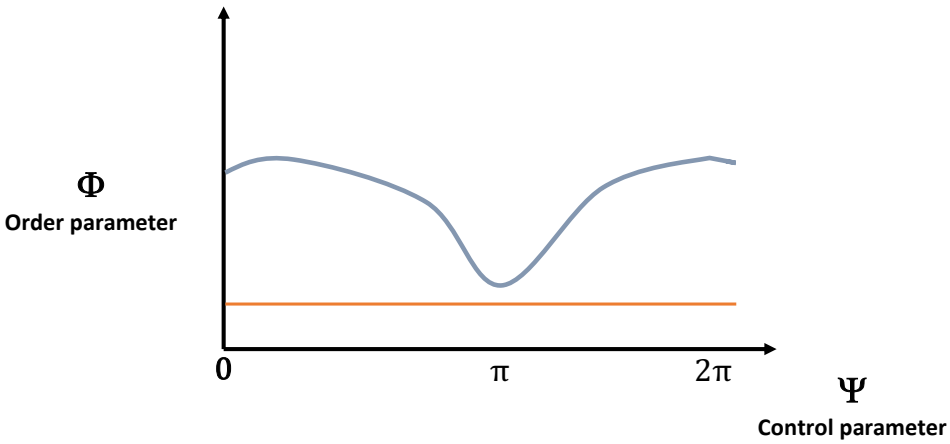


Figure 1. 3. The attractor landscape. Conceptual representation of the attractor points' layout, as returned by our experimental paradigm. The influence of attractor points is inferred by changes in the *order parameter* of the system (Φ) as the dyad transitions over the space of coordinative states, via manipulation of the *control parameter* (Ψ). The exploration of the space is guided by the drifting metronomes' cycle. The order parameter expresses the degree of coordination in the system and is quantified in our studies via a collective measure which will be presented in detail later on. At this point, it is sufficient to highlight that higher scores on this measure correspond to a higher degree of coordination in the dyad. Two toy examples are depicted to illustrate the concept. The blue line represents a scenario where the in-phase point, visited at the extremes of the cycle, operates as the only attractor in the space, resulting in stronger coordination around these regions. The orange line, on the other hand, represents the ideal case of two participants perfectly synchronizing with the assigned metronome, while showing no mutual influence. The latter case corresponds to the *ground-truth* in the context of this paradigm, providing a baseline of the expected coordination in absence of coupling.

Research questions and overview

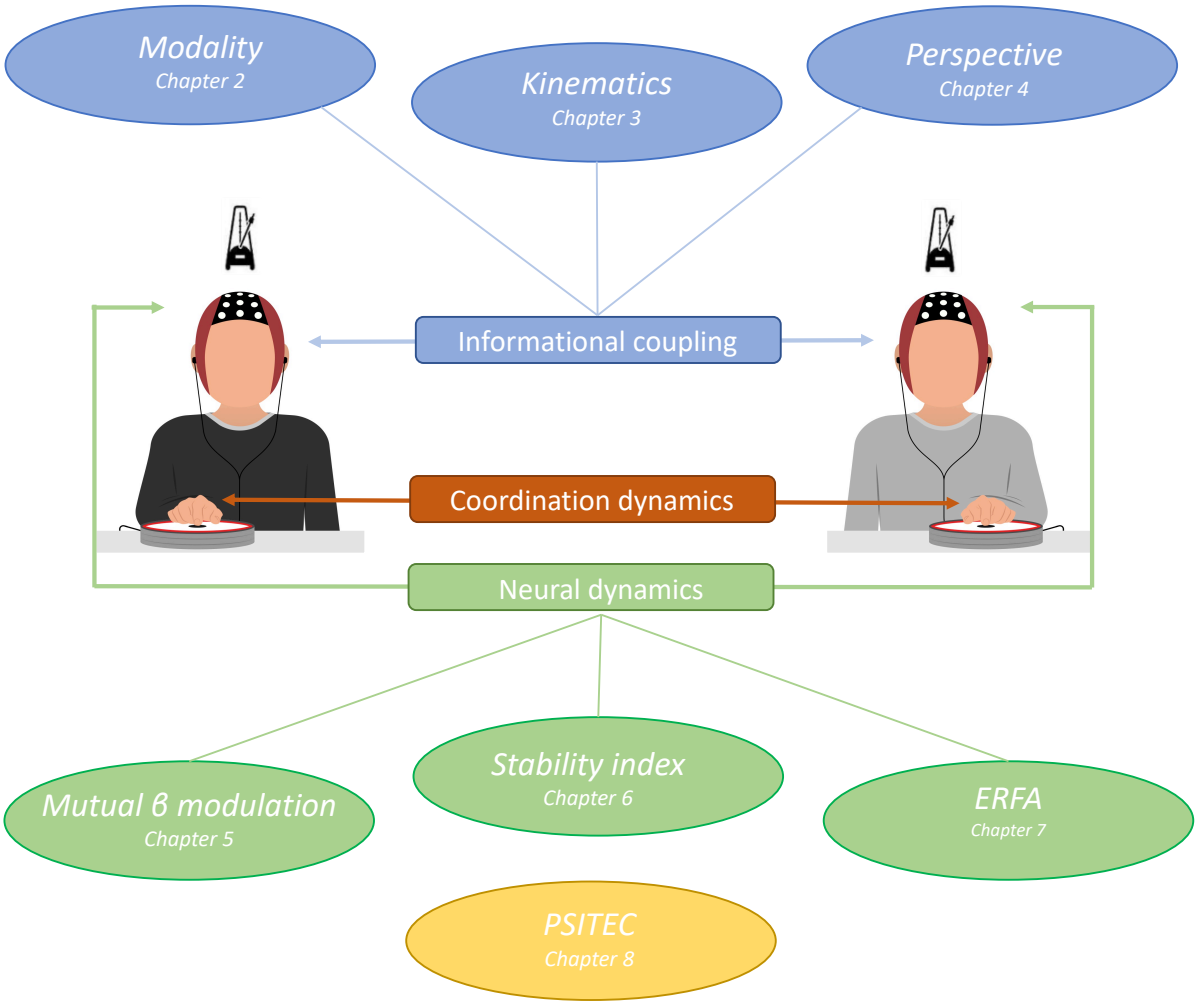


Figure 1. 4. Chapters overview. Block diagram of the topics covered in each chapter, connected to the major construct under investigation. The blue blocks deal with different aspects of *informational coupling*, while the green blocks focus on *neural dynamics*. Both constructs are foundational to the *coordination dynamics* of dyadic interactions. The yellow block represents a methodological contribution to the research program of *PSITEC* laboratory, based on the analysis techniques developed throughout our studies. The illustration highlights the cohesive nature of the dissertation, built on a unified methodological framework and rooted in its core theoretical concepts.

In *Chapter 1*, we introduced this dissertation by drawing out key theoretical concepts and articulating their interconnections, focusing on the fundamental aspects of rhythm and its prevalence in nature and biological systems. We examined the coordination dynamics that govern these interactions, particularly in relation to the control mechanisms regulating complex organisms such as humans. The crucial role of informational structures was emphasized, as well as the multifaceted challenge of understanding how separate individuals can achieve stable coordinated states through perceptual coupling. We situated our work within the context of major theories concerning human rhythmic interactions, inside and outside the social domain, and within the ongoing debates about the nature of the underlying mechanisms. Finally, we offered an encompassing perspective on the methodological concepts that guide our empirical studies, underscoring the shared principles that unify this body of work. It is now time to explicitly address its overarching main research questions.

Main research questions:

What are the effects of informational coupling on interpersonal coordination dynamics?

What are the neural dynamics underpinning coupled rhythmic behaviors?

As discussed, informational coupling in dyadic interactions is a multidimensional construct. We point out that while the scientific literature broadly recognizes its relevance, it has not engaged in a systematic empirical investigation of its components. Therefore, we lack a comprehensive picture of these dimensions and their specific influence on the dynamics of dyadic interactions. The main contribution of the present work consists of filling that gap, developing a research project aimed at building this fundamental knowledge. Starting from a theoretical reflection on the construct, we operationalized it by breaking it down into its constitutional elements. By doing so, we tackled more specific questions in a systematic and progressive approach.

Chapters 2, 3, and 4 consist of a series of studies where different dimensions of *informational coupling* were investigated. Starting from the most fundamental one and building up in complexity, we manipulated them as independent variables to assess their specific effect on the *coordination dynamics* of the interaction. Alongside the behavioral aspects, a major element of our investigation concerns the *neural dynamics* underpinning dyadic interaction. Specifically, Chapters 5, 6, and 7 are dedicated to identify the mechanisms deployed by the brain to couple its internal rhythms to external rhythms,

tracking their dynamics and enabling overt synchronization behavior. Chapter 8 consists of a separate methodological contribution in the context of the research performed at *PSITEC* laboratory, which builds on the principles of the dissertation to facilitate this line of research in normal and pathological aging.

The structure of the dissertation is presented in *Figure 1.4*. The block diagram indicates the subject of every chapter, and how they connect to the major constructs under investigation. We conclude this section by listing the specific research questions addressed in every study, alongside a brief introduction to the content of each chapter.

Question *Chapter 2:* *are attractor dynamics dependent on sensory modality?*

In *Chapter 2*, the first of a series of three studies, we investigate the foundation of perceptual coupling and its impact on the temporal coordination of individuals' movements. The study focuses on the most fundamental dimension, namely which sensory *modality* mediates an interaction. Visual and auditory couplings are manipulated in a factorial design, to assess their differential contributions to the coordination dynamics within a dyadic system. The drifting metronomes paradigm for dyadic entrainment is presented here for the first time.

Question *Chapter 3:* *are attractor dynamics dependent on the access to kinematic information?*

In *Chapter 3*, we make a distinction between the nature of the information available in the interaction, and the constraints imposed by sensory modalities in accessing such information. Moving from the observation that vision has preferential access to the continuous *kinematic* profile of an action (spatial specialization), whereas audition typically samples discrete events with high temporal accuracy (temporal specialization), we expand on the previous study by balancing the access to continuous and discrete information across modalities. We expect that differences between attractor dynamics across modalities would cancel out, provided equal access to the continuous kinematic profile of the partner's movement.

Question *Chapter 4:* *are attractor dynamics modulated by embodied perspective taking?*

In *Chapter 4*, we focus exclusively on the visual modality, examining the dimension of spatial processing in interpersonal coordination. To do this, we deploy virtual reality (VR) technology to manipulate the visual *perspective* (1st and 2nd person) on the action of the partner, under the hypothesis that perceiving movements from a 1st person perspective would enhance interpersonal synchronization, potentially mediated by the embodiment of the partner's effector.

Question Chapter 5: *how are beta oscillations involved in interpersonal synchrony?*

In *Chapter 5*, we enter the part of the dissertation concerning the neural dynamics underpinning rhythmic interactions in the social domain. In analyzing the hyperscanning EEG dataset recorded during the first dyadic experiment (see Chapter 2), we choose to approach the investigation from a new angle. Rather than focusing on interbrain measures of connectivity across brains of interacting participants (Zamm et al., 2023), we explicitly model the mediation of the physical effectors on their neural activity, via perceptual coupling. This approach allows us to examine the *modulation* of oscillatory dynamics as a function of reciprocal movement cycles, showing how the interaction between behavioral and neural oscillators is a potential underpinning of interpersonal synchronization.

Question *Chapter 6:* *is the stability of entrained oscillations an index for auditory-motor coupling?*

In *Chapter 6*, we leave the social domain to specifically focus on neural entrainment and its relationship to overt sensorimotor synchronization to auditory rhythmic stimuli. The work is based on an EEG dataset recorded during the first dyadic experiment, in a condition of finger-tapping to isochronous auditory metronome in absence of interpersonal coupling (see Chapter 2). We develop a *stability index* for neural entrainment based on the fluctuations of instantaneous frequency of components attuned to the rhythmic stimulus. We further test its correlations with behavioral outcome measure of synchronization consistency and accuracy, to assess its validity as a neural outcome measure of auditory-motor coupling.

Question *Chapter 7:* *does neural entrainment underpin sensorimotor synchronization to dynamic rhythmic stimuli?*

In *Chapter 7*, we build on the correlational evidence of Chapter 6, and develop a more sophisticated methodology to overcome some limitations and gain deeper insight on the dynamics of neural entrainment. Specifically, we propose *event-related frequency adjustment* (*ERFA*) as a paradigm to induce and measure neural entrainment in human participants, optimized for multivariate EEG datasets. Whereas our previous study developed an analysis pipeline on isochronous stimulation, here we apply phase and tempo perturbations to induce synchronization errors during a finger-tapping task. This allows us to zoom into adaptive changes in instantaneous frequency of entrained oscillatory components during error correction time windows.

Question *Chapter 8:* *how can we facilitate the investigation on SMS in healthy and pathological aging?*

In *Chapter 8*, we bring together the methodological concepts presented throughout the dissertation and develop analysis pipelines for two large datasets collected at *PSITEC* Laboratory (University of Lille, FR). Both studies investigate rhythmic musical interaction in healthy and pathological aging at different severity levels of cognitive disorder. The first dataset consists of force plate data recorded during a dyadic rhythmic task with a musical therapist (Ghilain et al., 2020), from which we compute three measures of varying complexity: 1) general quantity of motion, based on the body sway away from the center of gravity; 2) complex measures for temporal structure in these spatial patterns; 3) measures of coupling between the participants and the musician. The second dataset consists of finger-tapping data in a synchronization task to isochronous and tempo-changing auditory stimuli (von Schnehen et al., *in preparation*). The methodology for these behavioral data is based on Chapters 6 and 7. For both studies, we have to deal with major constraints due to the vulnerable target population in task design and data acquisition, which implies tailoring the analyses to suit our objectives and maximize our inferences in spite of limitations.

In *Chapter 9*, we draw general conclusions from the present work and address future directions to expand on the presented line of research.

2

Modality-specific attractor dynamics in dyadic entrainment

Mattia Rosso [A,B], Pieter-Jan Maes [A], Marc Leman [A]

[A.] IPEM Institute for Psychoacoustics and Electronic Music. Ghent University - Department of Art, Music and Theatre sciences.

[B.] Université de Lille, ULR 4072 – PSITEC – Psychologie: Interactions, Temps, Emotions, Cognition. Lille, France.

Rosso, M., Maes, P. J., & Leman, M. (2021). Modality-specific attractor dynamics in dyadic entrainment. Scientific Reports, 11(1), 1-13.

Introduction

Humans are in remarkable ways influenced by what they hear and see, especially when they perceive other people moving. The exposure to rhythmic behaviors attracts individuals to fall in sync with one another, and the phenomenon can be observed at the level of dyads, small-groups and large crowds (Zhang, Kelso, & Tognoli, 2018). For instance, when two people walk or dance together, they show a natural tendency to synchronize their steps (Miyata et al., 2017; Nessler & Gilliland, 2010; van Ulzen, Lamoth, Daffertshofer, Semin, & Beek, 2008). When an audience applauds at the end of the concert, the auditory scene goes through intermittent periods of synchronized clapping (Néda et al., 2000). When marching bands cross each other on the street, it requires conscious effort for them to avoid getting rhythmically entrained (Clayton, Dueck, & Leante, 2013).

Rhythmic joint coordination is ubiquitous in daily-life human activities, from the simplest form of unintentional synchronization (Richard C. Schmidt & Richardson, 2008) to complex tasks such as musical performance (P. Keller, 2014). The way different rhythms organize into orderly coordinated patterns over time implies co-regulated actions in response to rhythms produced by other agents. Characteristic for this co-regulation is that action patterns are often spontaneously attracted towards stable synchronization, via a process called *entrainment* (Phillips-Silver et al., 2010). *Coordination dynamics* has been proposed as a theoretical framework to understand the generic organizational principles that underlie the coordination of coupled rhythmic units (de Guzman & Tognoli, 2014; J. A. S. Kelso, 1995; Phillips-Silver et al., 2010; Richard C. Schmidt & Richardson, 2008; Schöner, 1989). Central to the theory is that the coordination of multiple units over time is the result of a dynamic balance between keeping the own rhythm (*competition*) and moving together on the collective level (*cooperation*) (Scott Kelso, Engstrom, & Engstrom, 2006). Such idea resonates in earlier work by von Holst(Von Holst, 1973), who termed the 'maintenance effect' versus the 'magnet effect' in his study of coordination in animal behavior. When intrinsic and external rhythms are incongruent, rhythmic units engage in a dynamic competition-cooperation process.

When it comes to rhythmic interactions within a dyad, this dynamic balance can be influenced by manifold variables, such as preferred tempo (R. C. Schmidt & Turvey, 1994), empathy (Novembre, Mitsopoulos, & Keller, 2019) and social factors (Tschacher, Rees, & Ramseyer, 2014). However, among all possible sources of variability, informational coupling between individuals stands out as the most fundamental condition for interpersonal coordination. The essential minimal requirement for human interactions to

occur is that individuals mutually exchange information through one or more sensory channels. Although experimental interactive scenarios necessarily imply a determined sensory modality (Ole A. Heggli et al., 2019; Konvalinka et al., 2010; R. C. Schmidt et al., 1990) and possibly cross-modal interactions (Miyata et al., 2017), the role of competing sensory modalities remains overlooked in the literature. To this date, whether and how coordination dynamics depend on the nature of available perceptual information requires further systematic investigation.

The aim of the present study is to develop a dynamic experimental paradigm, and apply it for studying modality dependencies of competition-cooperation dynamics in dyadic entrainment. The paradigm involves a unimanual tapping task, where two individuals are exposed to minimally different tempi and reciprocally coupled in visual and auditory modalities across experimental conditions. The procedure implies a tension between competition and cooperation processes in the dyadic interaction, and is meant to return an empirical layout of its dynamics across sensory modalities.

The core idea underlying the proposed experimental paradigm is to induce a competition between the intrinsic rhythm of an individual's actions and the rhythm produced by another person. During a unimanual tapping task, two partners are exposed to two metronomes in different modalities (visual and auditory). The metronomes assigned to the partners slightly differ in frequency, resulting in a gradual increasing (0 to π rad) and decreasing (π to 0 rad) of absolute relative phase over multiple cycles. Accordingly, for each person, a corresponding cyclical de-phasing of the assigned metronome and partner's metronome unfolds over time. With this simple expedient, we implemented a completely predictable system of two oscillators which deterministically revisits the same states throughout 10 consecutive cycles. An audio file showing metronomes' behavior in available in the *Supplementary materials.*

When the partners are informationally coupled, they are exposed to two competing rhythms. The first one is the *intended* rhythm: participants are explicitly instructed to synchronize to their assigned metronome. The second rhythm is the *coupling* rhythm, which is produced by the partner and perceived in another sensory modality. Thus, when the members of the dyad follow their own auditory metronome, they are exposed to visual information from the partner (visual coupling). Alternatively, when the members of the dyad follow their own visual metronome, they are exposed to auditory information from the partner (auditory coupling). The same task performed in absence of informational coupling provides the control conditions for each modality, and a baseline to assess the impact of coupling on spontaneous dyadic entrainment. In presence of informational

coupling, a cross-modal incongruency takes place between intended and coupling rhythms. The experimental design is schematized and described in *Figure 2.1*.

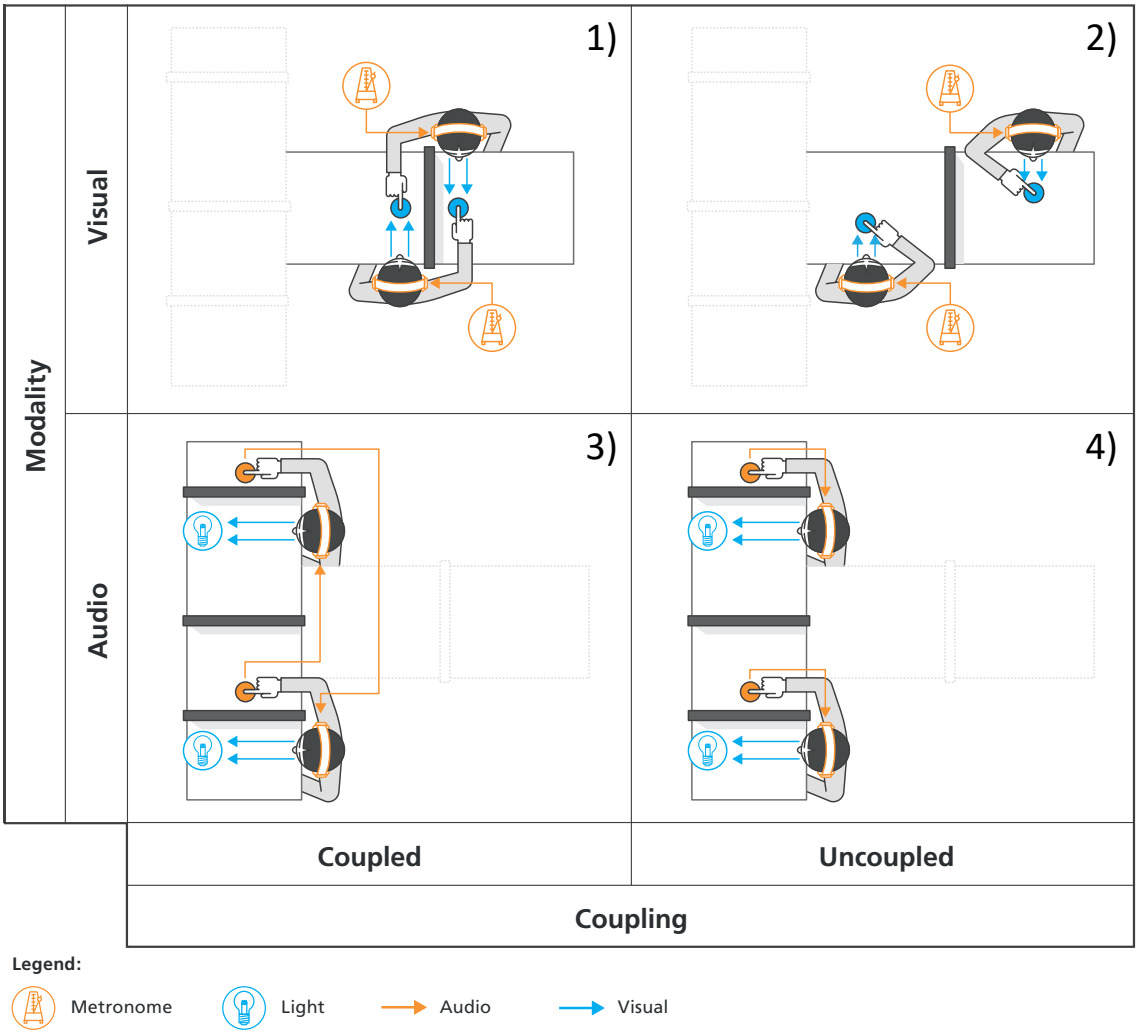


Figure 2. 1. Experimental design. The study was designed in a Modality (Visual, Auditory) x Coupling (Coupled, Uncoupled) factorial structure, resulting in the following conditions. ***1.Visually Coupled.*** Participants tapped along with an auditory metronome, while looking at the partner's hand tapping. The view of their own hand was hidden by a screen placed on the table. They were explicitly asked to neglect the partner's movements and focus on following their own metronome. ***2.Visually Uncoupled (control).*** Participants tapped along with an auditory metronome, while looking at their own hand. The view of the partner's hand is hidden by a screen placed on the table. ***3.Auditorily Coupled.*** Participants tapped along with a flickering LED, while hearing the sonification of the partner's tapping. The view of all the hands was hidden by screens placed on the table. They were informed that the sounds they would hear were produced by the partner, and were explicitly asked to neglect them and focus on following their own metronome. ***4.Auditorily Uncoupled (control).*** Participants tapped along with a flickering LED, while hearing the sonification of their own tapping. The view of all the hands was hidden with screens placed on the table. The present figure was realized by Kevin Smink, and published with his consent.

In the framework of coordination dynamics, the concept of *relative phase* (φ) functions as a collective variable to describe co-regulation of timing. When competition dominates the interaction, rhythmic units may fully keep their own intrinsic frequency, whereas they may be fully attracted towards each other as cooperation prevails. The former case would typically result in an evenly spread distribution across φ (0 to 2π radians). In the latter case, the distribution would exhibit peaks over the so-called *attractor* regions, where rhythmic units are locked in temporally stable state. In human behavior, the intrinsic dynamics of the collective variable φ are well characterized by a layout of attractive and repelling fixed points, wherein in-phase (0 radians) and anti-phase (π radians) are the dominant stable modes of coordination. The Haken-Kelso-Bunz (HKB) model mathematically describes how phase transitions from one spatiotemporal pattern to another take place when the system is pushed beyond its equilibrium state (Haken et al., 1985). The model points at the in-phase mode as the point where the system ultimately tends to reach stability.

Crucially, the behavior of the 'drifting metronomes' recursively guided the dyads through the whole range of φ values – and over its respective attractor regions. Based on a joint recurrence analysis of the system's dynamics (Marwan, Carmen Romano, Thiel, & Kurths, 2007), we computed a recurrence score as a metric of dyadic coupling strength and tracked its evolution over time (for a visual representation of the processing pipeline, see *Figures 2.2* & *2.3*). By these means, we obtained an empirical attractor layout descriptive of the competition-cooperation dynamics within the dyads. The cooperation process (i.e., entrainment to the partner's coupling rhythm) was expected to dominate the conflict around in-phase (0 radians) and anti-phase (π radians) points, leading to significantly higher recurrence score in presence of informational coupling and a significant modulation around attractor regions.

Given the time-varying nature of the recurrence score as a response variable, we needed a method to model its local variations over the course of the metronomes' cycles. We opted then for a statistical model which allowed us to assess the effects of our experimental manipulations over the entire unfolding of the interaction, and to avoid segmentation and multiple comparison while respecting the temporal structure of the data (Demos & Chaffin, 2017; Demos, Chaffin, & Logan, 2018). In addition to the inferences made possible by linear models based on intercepts, such as ANOVAs, our solution allowed to assess the effect of categorical predictors on the temporal profile of the response variable. In order to complement our observations at the dyadic level, we used the same framework to analyze the evolution of individual rhythmic behaviors over the course of the task.

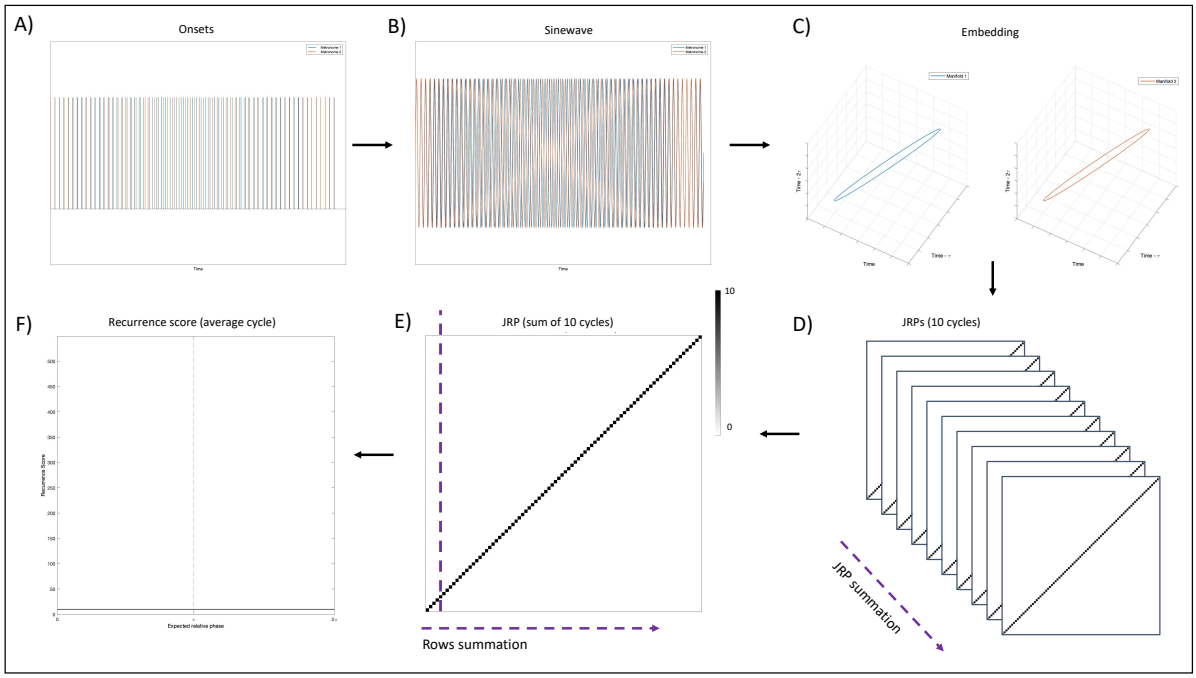


Figure 2. 2. Analysis pipeline. The procedure is here presented as applied to the timeseries generated by the metronomes' onsets, showing dyadic behavior under the ideal condition of perfect compliance with the task. A) Starting from discrete onsets, B) we interpolated each timeseries with a sine function to model the behavior as a system of coupled oscillators (Ole Adrian Heggli et al., 2019) and C) separately embed them in their respective phase-spaces for JRPs computation (Marwan et al., 2007). From the representation of one full cycle of the metronomes' behavior, it is evident how the two oscillators smoothly de-phase and get back in-phase at the end of the cycle. D) JRPs were computed of each individual cycle and E) summed up, returning a 2-D matrix of the density of recurrence points over the trial structure. Finally, by looping over the column and summing all the rows of the matrix, F) we compressed the representation into a 1-D timeseries of the total count of recurrences over the cycle structure. The presence of diagonal lines dominating the JRPs is exactly the configuration expected by a deterministic system (Marwan et al., 2007), and when compressed in the timeseries format result in a flat horizontal line with small periodic fluctuations (due to the down-sampling performed on the sinewaves prior to embedding, which was necessary to make JRPs calculations computationally feasible). Applying the pipeline to human dyads allowed us to map the local variations of their coupling strength over the expected relative phase, resulting in the picture of the 'attractor landscape' (Tuller & Kelso, 1989).

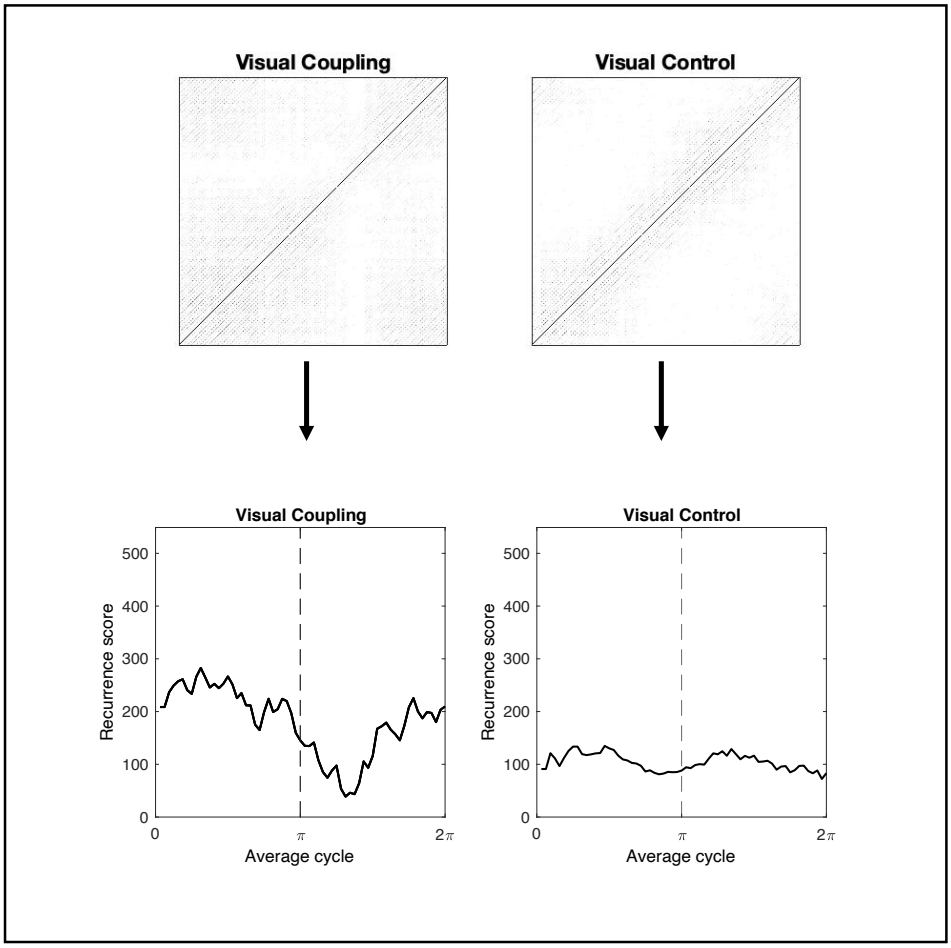


Figure 2. 3. Recurrence score over average cycle. The sum of 10 JRPs from Visual Coupling and Visual Control conditions are compressed to a 1-D recurrence score timeseries. It is evident how visual coupling modulates the density of recurrence point and its distribution over the average cycle in this representative dyad (#5). Compared to the control condition, the recurrence score was higher on average and exhibited local maxima at the extremities of the cycle (i.e., in proximity of the in-phase attractor point).

Results

Growth curve analysis (Mirman, 2017) was used to analyze the evolution of recurrence score over the course of the expected relative phase between partners (i.e., from 0 to 2π). The curves were calculated as the average of the 10 consecutive metronomes cycles in each experimental condition, and modeled with 2nd order orthogonal polynomials and fixed effects of Coupling and Modality on the polynomial terms. The uncoupled (control) conditions were treated as a baseline for contrasting the levels of the Coupling factor, and the audio conditions were treated as a baseline for the Modality factor.

In the context of orthogonal polynomials, a flat line can be considered a pure intercept and a 'zero-order' polynomial, in the sense that it exhibits zero changes in any direction. If the recurrence score was time-invariant, it would appear as a flat line indicating complete dominance of the competition process between rhythmic units, which would pursue their individual intended rhythms. On the other hand, significant changes of direction would indicate that the interaction is systematically modulated by the temporal structure of the task: around attractor points, the cooperation process would take over and the rhythmic units would be attracted towards the coupling rhythm. *Figure 2.4* shows how orthogonal polynomials are suited for modelling local variations of the response variable around attractor points.

Orthogonal polynomials in the model were limited to the 2nd order: based on our hypotheses, we expected quadratic and quartic terms to capture the effects of in-phase and anti-phase attractor points. However, following data inspection and analysis of residuals, we concluded that the quadratic term alone was enough to capture the relevant effect of our experimental manipulation. *Figure 2.5* shows the data with descriptive statistics across experimental conditions, and a comparison between 2nd order and 4th order models. Arguably, the better fit of the quartic term comes with a major cost in terms of model parsimony and interpretation. Our model also included random effects of Dyads on all polynomial terms, and their interactions with the factors: the random effects structure was maximized in order to minimize false alarm rates without substantial loss of power (Barr, Levy, Scheepers, & Tily, 2013).

As expected, we found a significant main effect of Coupling (*Estimate* = 42.809, *SE* = 18.142, p = 0.018) indicating an overall increase of the recurrence score in presence of informational coupling as compared to the uncoupled control conditions, independently from the sensory modality. We also report a significant interaction effect of Coupling and Modality on the orthogonal polynomial terms of Time (Linear: *Estimate* = -84.579, *SE* = 30.867, p = 0.006; Quadratic: *Estimate* = 190.320, *SE* = 36.973, p < 0.001), meaning that

the parameter estimates in coupled conditions significantly differ across visual and auditory modalities. Given the observed major variability in the Auditory level and following inspection of the residuals, we fitted the same model on the natural log-transformed response variable for it provided a considerably better fit, and still found significant main effect of Coupling (*Estimate* = 0.237, *SE* = 0.117, *p* = 0.042) and interaction effects on the polynomial terms of Time (Linear: *Estimate* = -0.534, *SE* = 0.222, *p* = 0.016; Quadratic: *Estimate* = 1.257, *SE* = 0.299, *p* < 0.001). *Table 2.1* shows the fixed effects parameter estimates and their standard errors for log-transformed recurrence score, along with p-values estimated using the normal approximation for the t-values.

The same model was then fitted to the phase error of the individual participants. The response variable was computed as the absolute phase difference between the finger-tapping and the metronomes timeseries, to quantify individual synchronization behavior. We found a significant main effect of Modality (*Estimate* = -0.230, *SE* = 0.080, *p* = 0.004), indicating that the assignment to visual metronomes resulted in higher phase error. Furthermore, the quasi-significant main effect of Coupling (*Estimate* = 0.156, *SE* = 0.080, *p* = 0.051) confirmed that informational coupling with the partner negatively affected the synchronization with the assigned metronome. When looking at the evolution of the phase error over time, the interaction effect of Coupling and Modality on the quadratic term of Time (*Estimate* = 1.058, *SE* = 0.373, *p* = 0.004) revealed that the temporal profile across conditions is coherent with the one observed for the recurrence score (see *Figure 2.6*).

In the context of growth curve analysis (Mirman, 2017), the parameter estimates provide a measure of effect size of straightforward interpretation for linear and non-linear changes over time as long as the order is not too high: above the 4th order, results become hardly interpretable and the risk of over-fitting the data increases. With the interaction effect of Coupling and Modality factors on the polynomial terms, we could quantify modality-specific effects of informational coupling on the evolution of the recurrence score and the individual synchronization performances. The results strongly support our predictions of spontaneous dyadic entrainment in presence of informational coupling and, crucially, show that its temporal dynamics are modality-dependent.

	Recurrence score (N = 14)		
Predictors	*Estimates*	*SE*	*p*
Time	-0.037	0.113	0.739
$Time^2$	-0.360 *	0.150	**0.017**
Modality	0.155	0.117	0.184
Coupling	0.237 *	0.117	**0.042**
Time:Modality	-0.021	0.157	0.893
$Time^2$:Modality	0.269	0.212	0.202
Time:Coupling	0.011	0.157	0.942
$Time^2$:Coupling	0.055	0.212	0.793
Modality:Coupling	0.118	0.165	0.475
Time:Modality:Coupling	-0.534 *	0.222	**0.016**
$Time^2$:Modality:Coupling	1.257 ***	0.299	**< 0.001**
	** p < 0.05*	*** p < 0.01*	**** p < 0.001*

Table 2. 1. 1. Model's summary (dyadic behavior). All the significant effects are marked with an asterisk, and the associated p-values are highlighted in bold. We reported a significant main effect of Modality and significant 3-way interactions with linear and quadratic functions of Time. In the framework of orthogonal polynomials, coefficient estimates represent a measure of effect size (Mirman, 2017): the intercept represents the average increase in the response variable contrasted to the baseline (i.e., the Uncoupled level), the slope of the linear term represents the change in the response variable per unit of time, and the quadratic term represents the steepness of the parabolic inflection. For the recurrence score, the natural exponential of the coefficient estimates should be used for interpretation purposes, given that the response variable was transformed to the log-scale. The significant correlation with the quadratic function of time was not discussed in the text, for it does not have relevance in the context of the experimental design.

	Phase error (N = 28)		
Predictors	*Estimates*	*SE*	*p*
Time	-0.177	0.165	0.283
$Time^2$	-0.172	0.190	0.364
Modality	-0.230 **	0.080	**0.004**
Coupling	0.156	0.080	**0.051**
Time:Modality	0.138	0.218	0.526
$Time^2$:Modality	0.027	0.264	0.917
Time:Coupling	0.293	0.218	0.179
$Time^2$:Coupling	0.191	0.264	0.467
Modality:Coupling	0.183	0.113	0.106
Time:Modality:Coupling	-0.404	0.309	0.190
$Time^2$:Modality:Coupling	1.058 **	0.373	**0.004**
	** p < 0.05*	*** p < 0.01*	**** p < 0.001*

Table 2. 1. 2. Model's summary (individual behavior)

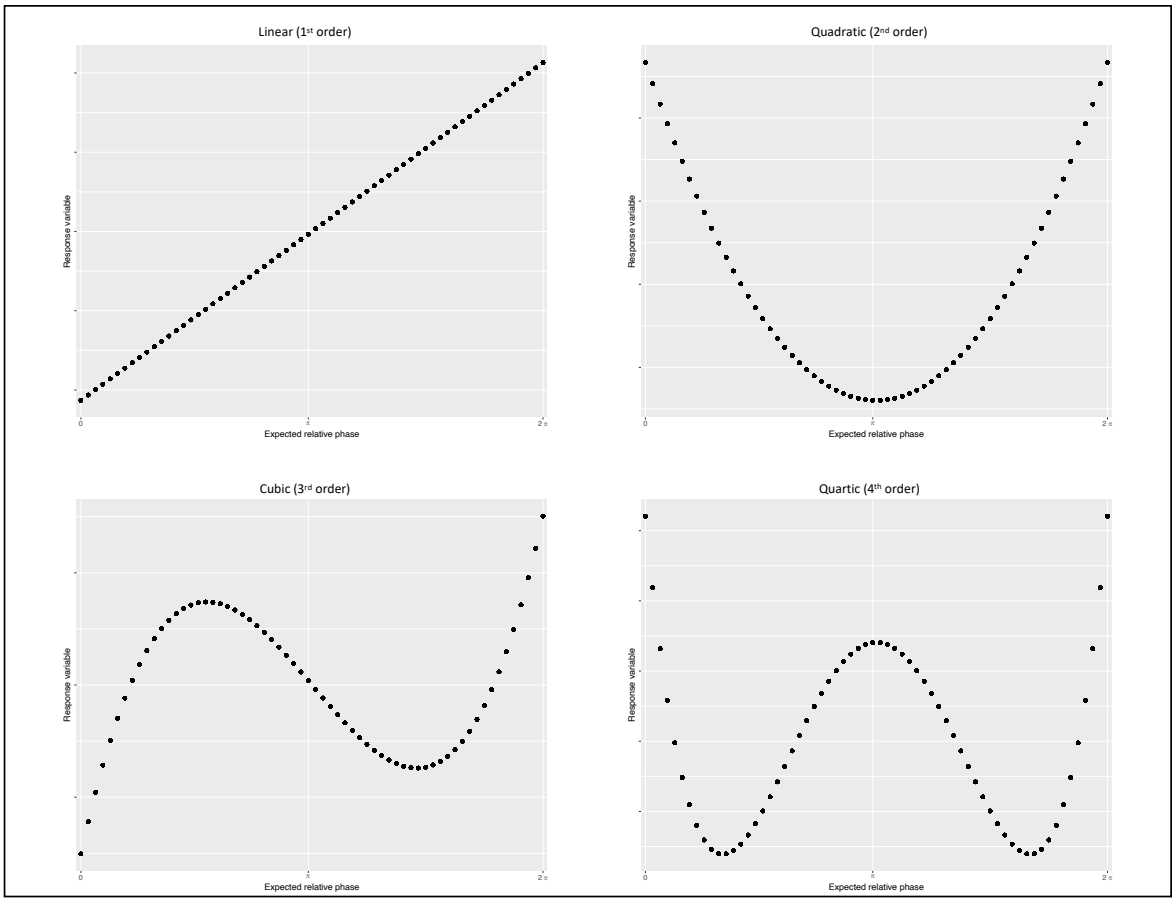


Figure 2. 4. Orthogonal polynomials. The time-varying dimension of a response variable over the metronomes cycle can be modelled with higher-order polynomials. In our model, intercepts correspond to the overall average of the recurrence score: the main effects of categorical predictors indicate global differences across experimental conditions independently from the temporal profile of the response variable. On the other hand, effects of categorical predictors on the polynomial terms indicate the local effect of attractor points on the x-axis. For the linear component (1st order), the parameter estimate corresponds to the slope of the line; for the quadratic component (2nd order), it corresponds to the parabolic curvature; for the cubic and quartic components (3rd and 4th order), it corresponds to the sharpness of the peaks on the inflexion points. The order of the polynomial is ideally chosen based on hypothesis and on the nature of the data, should be confirmed by the empirical data, and should be allow a straightforward interpretation of the effects (Mirman, 2017). In our case, orders from 2 to 4 seem suited candidates for modelling different possible attractor landscapes, given the symmetric measure of the cycle and inflection points around in-phase and anti-phase regions. A full model based on orthogonal polynomials includes all terms up to the chosen order.

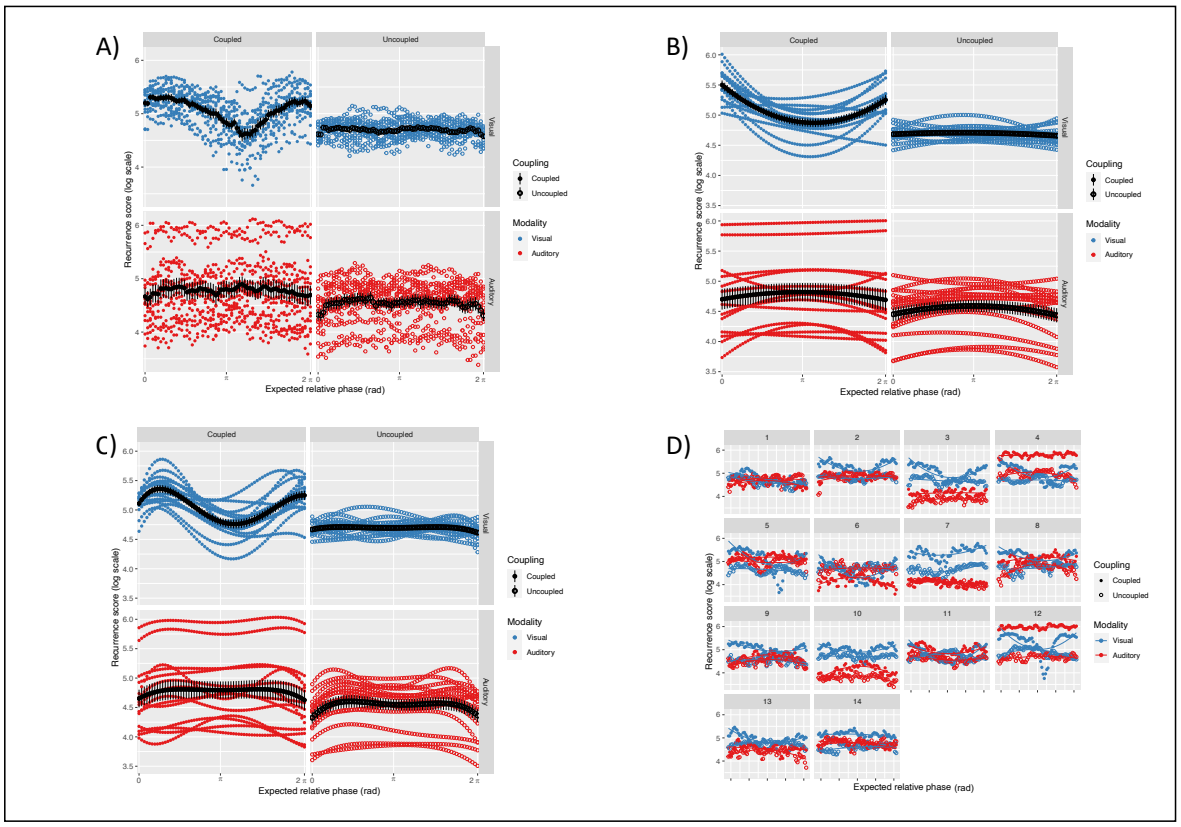


Figure 2. 5. Coupling strength as a function of the attractor landscape. *A) Recurrence score (on the log-scale) over the average of 10 cycles, for 14 dyads (colored dots)*. Grand-average and error bars (standard error of the mean, SEM) are represented in black. The reader can visually appreciate the consistency of the "seagull-shaped" pattern in condition of visual coupling and the flat line in the respective control condition. The condition of auditory coupling shows a recurrence score higher on average but no clear pattern over the average cycle, and overall higher variability. *B) Recurrence score predicted by our 2nd order polynomial model.* The inflection of the parabolic term and the negative slope of the linear term successfully capture the influence of the in-phase attractor points in the visually coupled condition. Arguably, this is the most parsimonious model for our data in the framework of growth curve analysis (Mirman, 2017). *C) Recurrence score predicted by our 4th order polynomial model.* The quartic fit manages to capture more fine-grained modulations, particularly the local maxima on the inflection points at the extremes of the cycle. The profile recalls the typical 'seagull-effect' originally reported in (Tuller & Kelso, 1989) and discussed in (J. A. S. Kelso, 1995) in the context of bimanual coordination dynamics. However, for the sake of interpretation of parameter estimates, we opted for the simplest possible model: given the absence of a local maximum on the anti-phase point, a single parabolic inflection point was sufficient to explain the effect of visual coupling. *D) Within-dyads comparison of the recurrence score over the average trial, across experimental conditions*. Dots represent the empirical data, whereas continuous lines represent the full model (up to the 2nd order polynomial) fitted to the data.

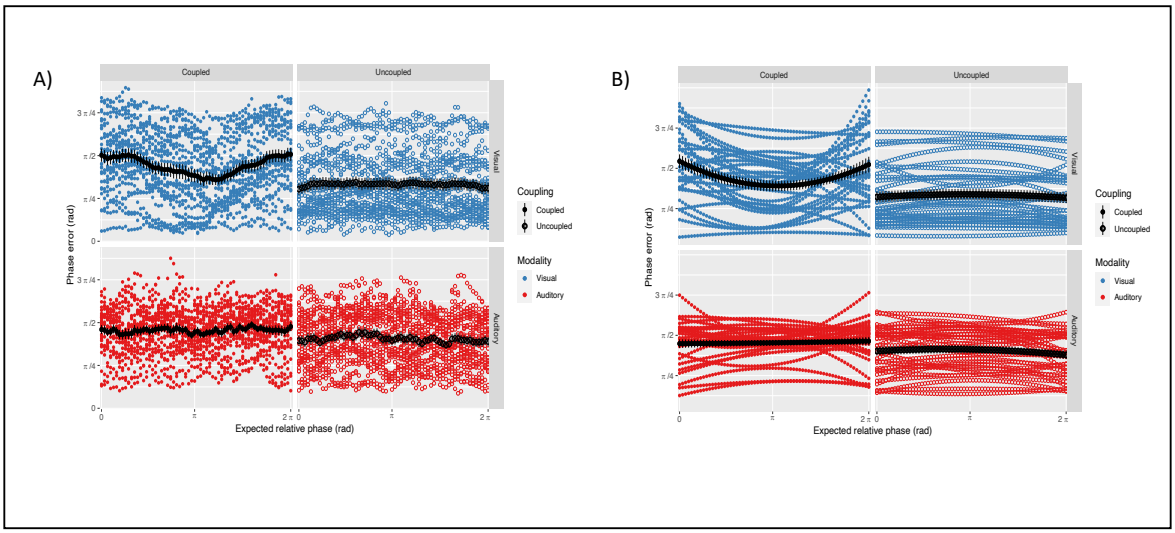


Figure 2. 6. Individual synchronization as a function of the attractor landscape. *A) Phase error over the average of 10 cycles, for 28 participants (colored dots).* Grand-average and error bars (SEM) are represented in black. The reader can visually appreciate how individual rhythmic behaviors track the course of the recurrence score, supporting the idea of a dynamic balance between cooperation and competition processes. As the coupling between partners dominates the interaction, the magnitude of the phase error increases accordingly. *B) Phase error predicted by the 2nd order polynomial model.* The "seagull-shaped" pattern observed in the visually coupled condition could be well modelled by the inflection of the parabolic term alone. Differently from the recurrence score, phase error exhibited a symmetric profile without significant interactions with the linear term.

Discussion

The present study investigated the dynamics of dyadic entrainment, and more specifically how they depend on the sensory modality mediating informational coupling. An experimental paradigm was designed to dynamically manipulate the expected relative-phase between participants, while informational coupling was manipulated in visual and auditory modalities across conditions. At the global level, we wanted to prove that when intended and coupling rhythms are at odds, a spontaneous co-regulation of timing occurs within the dyad at the expenses of the individual intended rhythms. At the local level, we wanted to infer the configuration of a dyadic *basin of attraction*, by modelling the dyadic cooperation process around attractor points (J. A. S. Kelso, 1995). Ultimately, it was possible to evidence a crucial modality-dependency of dyadic entrainment.

In the first place, we reported that the mere presence of informational coupling led to spontaneous entrainment in dyads engaged in a rhythmic task. As long as two partners perceived each other, their rhythms were 'attracted' towards each other. The effect was quantified as a significant increase of the average recurrence score, compared to the baseline provided by the de-coupled conditions. Remarkably, such manifestation of self-organizing behavior (J. A. S. Kelso, 1995; Tognoli et al., 2020) emerged in the context of a task where subjects were explicitly instructed to ignore each other's actions, pursuing independent uncoupled behaviors cued by their own reference. Even though the observation of spontaneous dyadic entrainment was reported in self-paced (Konvalinka et al., 2010; Lorås, Aune, Ingvaldsen, & Pedersen, 2019), cued (Kimura, Ogata, & Miyake, 2020; Okano, Shinya, & Kudo, 2017) and joint improvisation tasks (Noy, Dekel, & Alon, 2011; Noy, Levit-Binun, & Golland, 2015; Varlet, Nozaradan, Nijhuis, & Keller, 2020), our procedure explicitly implemented rhythmic incongruency to investigate the competition-cooperation dynamic balance latent to the system's dynamics (Scott Kelso et al., 2006). When competition prevails, rhythmic units manage to keep the assigned (intended) rhythm, whereas when cooperation takes over, rhythmic units undergo a (spontaneous) co-regulation of timing and move together on a collective level of coordinated behavior. Overall, both visual and auditory coupling led to spontaneous dyadic entrainment despite the instruction of synchronizing with the assigned metronome. Moreover, the shift towards a cooperation process occurred at the expenses of the competing individual behaviors, as indicated by the significant increase of the average phase error in coupled conditions.

Beyond a global measure of coupling, the continuous de-phasing implemented in the drifting metronomes paradigm allowed us to quantify local variations of recurrence score as a function of the expected relative phase between partners. The balance between

cooperation and competition processes is expected to be dynamic, in the sense that it may vary over time and is constrained by the intrinsic dynamics of the system. Dyadic entrainment was expected to occur more strongly and more consistently around attractor regions, located around the in-phase and anti-phase points of the metronomes' cycles (Haken et al., 1985). With our approach, we achieved to return an empirical attractor landscape (Tuller & Kelso, 1989) and to contrast it across conditions of visual and auditory coupling. Our findings strongly point towards a modality-dependency of dyadic entrainment (see *Figure 2.5*), which in turn is reflected in the individual synchronization with the metronomes (see *Figure 2.6*). The interaction effect of Coupling and Modality on the linear and quadratic terms of Time statistically confirms that the temporal modulation of the recurrence score is an exclusive property of visually-mediated coupling, with a remarkable consistency across dyads. In such condition, the recurrence score exhibits maximum values around the in-phase point at the start of the metronomes cycle, gradually decreases as the dyad leaves the attractor zone, reaches its minimum passed the anti-phase point, and finally increases towards maximum values as it enters the in-phase zone at the end of the cycle. On the other hand, the individual phase error exhibits a more symmetrical parabolic modulation, with only interaction on the quadratic term of Time reaching significance.

Selectively for the recurrence score, we report here an effect of directionality in the visually coupled condition: dyads slowly de-couple as they leave the in-phase point, whereas they couple more rapidly as they approach the same point. The negative slope of the linear term indicates that dyads tend to be more coupled in the first portion of the cycle, namely before passing the anti-phase point. The observation is suggestive of a hysteresis effect, which characterizes transitions from one attractor to another in dynamical systems (J. A. S. Kelso, 1995). In interlimb and interpersonal coordination, non-linear transitions from anti-phase to in-phase coordinative modes are typically observed as the system is pushed towards instability (J. A. S. Kelso, Scholz, & Schöner, 1986; R. C. Schmidt et al., 1990; R. C. Schmidt & O'Brien, 1997). This can be experimentally induced in an experimental setting by increasing the frequency of the rhythmic units as control parameter. In our paradigm, we could appreciate that increasing the metronomes' relative phase away from the in-phase point, the attractor exerted a long-lasting 'pull-back' effect. This was followed by a more abrupt 'push-forward' when transitioning over the anti-phase point, as the expected relative phase decreased. Even though the metronomes' frequencies were not manipulated, their continuous de-phasing resulted in repeated transitions from in-phase to anti-phase regions and vice versa, as the dyad explored the attractor landscape.

In condition of auditory coupling, the recurrence score did not exhibit any significant variation over time: the measure was locally independent from the proximity of attractor points. Nevertheless, we still observed a significant global increase of recurrences compared to the baseline. Previous evidence shows that phase shifts in auditory sequences elicit a phase-correction response in participants instructed to tap in synchrony to isochronous sequence of flashing lights (Bruno H. Repp & Penel, 2002). Auditory-driven phase corrections are prone to occur even when shifts are highly irregular (Kato & Konishi, 2006) and at different relative phases values, with overall high interindividual variability (Bruno H. Repp & Penel, 2004). Similarly, in the context of our paradigm, sounds generated by the partner exerted attraction on the self-generated taps over the whole course of the metronomes' cycle, regardless of the expected relative-phase value.

Finally, both control conditions resulted in a flat horizontal line below the average of the respective coupled conditions, as expected. It is worth noting the higher variability in Audio conditions, arguably due to the task of synchronizing to visual metronomes. From lower recurrence score in the control condition, we can conclude that dyads performed globally worse in synchronizing to visual cues as compared to auditory cues. This is in line with previous evidence on modality-dependent synchronization skills (Comstock et al., 2018; Iversen, Patel, Nicodemus, & Emmorey, 2015), and backed by our concomitant observation of higher phase errors when visual metronomes were presented. Taken together, our findings support our hypothesis of modality-dependent dynamics, suggesting that auditorily-mediated entrainment is considerably less sensitive to the basin of attraction when compared to its visual counterpart.

Although interpersonal coordination can be described at the collective level as a self-organizing process (J. A. S. Kelso, 1995; Tognoli et al., 2020), humans ultimately entrain to each other via mutual adaptation of individual action-perception loops (Ole A. Heggli et al., 2019; Konvalinka et al., 2010). Therefore, an exhaustive account for dyadic entrainment should be capable of bridging the two levels, considering how coordination dynamics depend on the processing of information available in the interaction. Such link is foundational for translating dynamical control principles from intra- to inter-personal levels (R. C. Schmidt et al., 1990). Recently, predictive coding (K. Friston, 2005, 2010; K. Friston, Mattout, & Kilner, 2011) was put forward as a plausible theoretical account for coordination dynamics in dyadic behavior (Koban, Ramamoorthy, & Konvalinka, 2019). The theory states that the brain is constantly engaged in an optimization process based on Bayesian inference: information sampled from the environment is compared to prior evidence, and the optimization consists of minimizing prediction errors. In the context of interpersonal coordination, such a putative mechanism could underpin the overt manifestations of dyadic entrainment by engaging the action systems of the rhythmic

units involved. Let us consider the prediction error as the phase difference between executed and observed movements: as both units minimize the mismatch between the representations of observed and own motor behavior, the dyadic system tends towards a collective minimization of free energy in a stable attractor state (Koban et al., 2019). Such interpretation of the dyadic attractor is supported by our results in the context of visually-mediated coupling: as the relative phase between partners increases, the cooperation process between rhythmic units loses strength in favor of competition. In other terms, the tendency of the partners to lock in a stable coordinative state depends on the energy required to correct for the prediction error. As the error increases, the correction becomes more effortful and it becomes easier for the units to pursue independent behaviors.

Crucially, such explanation does not hold for auditorily-mediated coupling. As distinct sensory systems have their unique interface with motor and timing systems (Comstock et al., 2018; Hove, Fairhurst, Kotz, & Keller, 2013; Hove, Iversen, Zhang, & Repp, 2013), it is necessary to interpret the results in light of individual information processing. There is arguably a crucial difference in the nature of the perceptual information available to the partners across different modalities. When the coupling is visually mediated, a stream of kinematic information is continuously sampled from the partner's actions (James M. Kilner, Friston, & Frith, 2007), such that the predictive models can be constantly updated at every stage of movement execution. On the other hand, in case of auditory coupling the prediction of the partner's tap is solely based on temporal information about the previous taps, since partners do not have online access to the reciprocal kinematic information. In such condition, the cooperation process still dominates the interaction but no consistent pattern emerges in terms of dynamics as quantified by our method. We propose that the availability of kinematic information through informational coupling is a crucial discriminant that might explain the observed modality-dependence of dyadic entrainment. The idea that prior evidence is built differently across modalities depending on availability of sensory information is consistent with a predictive account for dyadic entrainment.

A final question arises spontaneous: how would the dyad behave if kinematic information was embedded in a sound envelope, and conveyed to the partners via the auditory channel? One limitation of the current work is that we cannot conclude whether dyadic entrainment is modality-dependent in the strict sense or rather dependent on availability of kinematic information. Given the central role attributed to kinematics in predictive accounts for action perception (James M. Kilner et al., 2007), it is of particular interest to investigate whether continuous sonification of movement parameters (Maes, Buhmann, & Leman, 2016) would support the same dynamics observed in presence of visual coupling. Previous work shows that presenting visual stimuli in a way that indicates

movement over time, e.g., apparent hand motion (Hove & Keller, 2010; Hove, Spivey, & Krumhansl, 2010) or a bouncing ball (Gan, Huang, Zhou, Qian, & Wu, 2015; Hove, Iversen, et al., 2013; Iversen et al., 2015), behavioral outcomes of predictive entrainment are improved by the continuous stream of information. We hypothesize that a movement sonification strategy based on this principle could translate the kinematic advantage into auditorily-mediated dyadic interactions, resembling the dynamics here described in condition of visual coupling.

The findings and the questions raised from the present study might have relevant implications for optimizing interventions aimed at rehabilitating, teaching or training sensorimotor functions. Fundamental knowledge about dyadic interactions is what will ultimately inform such optimization. The evidence from the present study points at the advantages of the physical presence of a teacher (or a therapist, or a trainer) for guiding rhythmic movements, possibly integrating kinematic information via sonification strategies (Maes et al., 2016).

In conclusion, we want to highlight that the drifting metronomes procedure is meant as a methodological contribution for the investigation of interpersonal coordination, in the hope that the scientific community can build upon it with new research questions and experimental designs. We point at its adoption with simultaneous dual-electroencephalography recordings (Liu et al., 2018), computational simulations (Ole Adrian Heggli et al., 2019) and replication on pathological populations (Dermody et al., 2016; Marsh et al., 2013; Varlet, Marin, Raffard, et al., 2012) as potential sources of insight into the fundamentals of dyadic interactions.

Methods

Participants. Twenty-eight right handed participants took part in the study (18 females, 10 males; mean age = 29.07 years, standard deviation = 5.73 years). They were randomly paired in fourteen (N = 14) gender-matched dyads, in order to control for any gender bias in the interaction. None of them had history of neurological, major medical or psychiatric disorders. All of them declared not to be professional musicians upon recruitment, although some of them had musical experience. With the exception of one dyad, all participants declared not to know the assigned partner from before the experiment. The experiment was approved by the Ethics Committee of Ghent University (Faculty of Arts and Philosophy) and informed written consent was obtained from each participant. A 15€ coupon was given to all participants as economic compensation for their time.

Experimental apparatus and procedure. The two partners were sitting across the same table, facing each other. Chairs were provided with an armrest in order to exclude any tactile or proprioceptive coupling due to the table's vibrations. Each partner was assigned to one pad and instructed to tap on it with the right index finger synchronizing with a metronome. As represented and described in *Figure 2.1,* the metronome was either an auditory cue or a flickering light, depending on the experimental condition. Each partner was cued with a different tempo (1.67Hz and 1.64Hz). Aligning the start of the two tracks, the relative phase between the metronomes started at 0º and steadily increased in regular steps of 5.6º. A full cycle took 39 seconds to be completed. 10 consecutive cycles were performed in each experimental condition. In conditions of informational coupling, participants were instructed to ignore the partner and tap along with the assigned metronome.

A M-Audio M-Track 8 soundcard was used to route independent audio channels to each participant via in-ear plugs. Ableton Live 10 was adopted as main interface for stimuli presentation, to sonify the participants' finger-taps and to route them in real-time. The same MIDI tracks were used to control the metronomes across conditions, by either triggering an audio sample or a flickering LED. Volume was adapted to every participant before the start of the experiment, and pink noise was regulated up to the point of suppressing any sound other than the auditory stimulation presented via Ableton. Participants were monitored on-line by means of a USB-camera, to make sure they complied with the instructions. No dyad was excluded from the analyses.

A Teensy 3.2 microcontroller was used as serial/MIDI hub in the setup. It was used to detect tapping onsets with 1 ms resolution, based on the analog input from strain gauge pressure sensors installed inside the pads. Every time a metronome onset was presented to a participant, a MIDI message was sent to the Teensy device to log the metronomes timeseries and to control the voltage of the LEDs when needed.

Simultaneous EEG recordings were performed from both partners of the dyads during the whole experiment, but such data were not presented in the present paper. Additional data were collected prior and during the experiment. Prior to the experiment, demographical data were collected; the *Edinburgh inventory* (Oldfield, 1971) was administered to assess the right handedness of the participants; the *Interpersonal Reactivity Index* (IRI) was administered as self-report of empathy and its subscales (Davis & Others, 1980); the preferred spontaneous tempo was calculated via 30 seconds of self-paced finger-tapping on a dedicated smartphone app (*www.beatsperminuteonline.com*). During the breaks between experimental conditions, all participants provided subjective self-reports on

different aspects of the task by expressing agreement on a scale from 1 ("Completely disagree") to 7 ("Completely agree") with a custom-made battery of 11 Likert items. Data collected from the questionnaires were not presented in the present paper.

Data Analysis

Pre-processing. Our raw data consisted of timestamps logged from the Teensy controller with 1ms resolution for 10 consecutive metronome cycles (390 seconds in total), and an associated ID for each partner and each metronome. As only form of cleaning, we removed onsets occurring <350ms from the previous one: false positives could occasionally be recorded when a participant pushed the pad for too long or accidentally laid the hand on it. The cleaned timeseries were then interpolated with a sine function at 1kHz sampling rate, providing an estimate of the oscillators' positions on its cycle with a temporal resolution of 1ms. Conceptually, the choice of sinusoidal interpolation was supported by recent work on modelling of systems of coupled oscillators in joint finger-tapping studies (Ole Adrian Heggli et al., 2019). Operationally, it guaranteed that all timeseries match in size without any loss of data, which was a requirement for the next steps of our analysis. Timeseries were finally down sampled by a factor of 40 to make the computation of RPs computationally feasible. Different orders of down sampling were tested to make sure that the results do not depend on this choice.

Phase-space reconstruction. The optimal parameters for the time-delayed embedding were computed for each participant, for the time course of each single metronomes cycle in all experimental conditions. The resulting mean value across all participants was applied to all individual instances. The reason for this approach is that in order to compare the rate of recurrences across conditions at the group level, the embedding procedure must be consistent across participants (e.g., see Afsar, Tirnakli, & Marwan, 2018, for an example of parameter selections in a factorial design). We first selected the delay *tau* of the timeseries (τ) as the first local minimum mutual information index (Fraser & Swinney, 1986) as a function of delay. This approach minimized the timeseries self-similarity, extracting nearly orthogonal components and preventing the attractor from folding over itself (Bradley & Kantz, 2015). The mean value of τ resulted to be 7. Next, we determined the number of embedding dimensions *m* with the method of false nearest neighbor (Rulkov, Sushchik, Tsimring, & Abarbanel, 1995): we progressively unfolded the time series into higher dimensions until the data points did not overlap spuriously, finding a mean *m* = 3. Finally, in order to determine a maximum threshold for counting two neighboring

points as recurrent, we selected a *radius* of 10% of the maximal phase-space diameter (Marwan et al., 2007).

Joint Recurrence Plots (JRPs). Individual recurrence plots were computed as follows:

$$R_{i,j}(\varepsilon) = \Theta(\varepsilon - \|x_i - x_j\|)$$

where ε is the neighborhood threshold, ‖ · ‖ is the Euclidean norm, and Θ is the Heaviside step function. A square matrix was returned from each shadow-manifold in the phase-space, containing 1s for all the instances where the distance ‖ · ‖ was smaller than the threshold ε, and 0s for the remaining elements. The distance was computed with the method of maximum norm. A joint recurrence plot (JRP) was computed for each dyad by overlapping the individual JRPs of the partners pair-wise, and keeping as 1 only the instances where both plots contained a recurrence. The computation of the JRPs was carried out with the *crp toolbox* for Matlab (Marwan et al., 2007).

The 10 trials (i.e., the metronome cycles) of each experimental condition were aggregated by summing the respective JRPs such that every cell of the 2-D matrix contained the count of times that a recurrence occured in the same point of the cycle. Finally, we looped over the columns of the matrix summing all the counts contained in the rows, obtaining a 1-D vector recurrence scores which represented a density measure of the instances of coupled behavior over the course of the cycle. The scale of the recurrence score depends on the size of the JRPs and in turn on the embedding procedure, which maked it necessary to set the same parameters for the whole sample. In order to stabilize our response variable and avoid over-sampling in view of our statistical model, the resulting timeseries were divided in 64 segments averaging the recurrence score. For the division, we chose the intervals determined by the steps of the slower metronome, as they provided an intrinsic regular subdivision of the experimental runs. All the steps presented so far were carried out in Matlab (see *Figure 2.2*, for a schematic representation). Our approach was preferred over the "windowed" version for JRQA, for the latter would low-pass filter our timeseries and make it impossible to interpret our results. The resulting phase-shift would be dependent on the choice of the window size, hence not reliable for detecting attractor points over the landscape.

Individual rhythmic behavior. For every participant, phase angle timeseries were computed by linearly interpolating the finger-tapping onsets as a rampwave at 1kHz sampling rate, and scaling it by 2π. The same procedure was repeated for the metronomes' onsets. Phase error was calculated as the absolute difference between the participants' and the assigned metronomes' timeseries, wrapped to π and averaged within 64 bins like the recurrence score timeseries. This measure was used as response variable to assess the individual rhythmic behavior of the participants, complementing our findings at the level of collective behavior.

Statistical models. The recurrence score was used as response variable in a mixed-effects model with Modality and Coupling as factors, and Time as a continuous predictor expressed with the indexes of the metronome's steps (from 1 to 64). Given the non-linear time course observed in the 'Visual Coupling' condition, we adopted the method of orthogonal polynomials (Mirman, 2017) including linear and quadratic functions of Time into our model (see *Figure 2.5*). Dyads and interactions between Dyads and factors were modelled as random effects (Barr et al., 2013) to account for the individual variability in synchronization skills and individual susceptibility to coupling across the experimental manipulations. The formula of the full model (up to the 2nd polynomial order) is the following:

$$Recurrence \sim (Time + Time^2) * Modality * Coupling + (Time + Time^2 \mid Dyad) + (Time + Time^2 \mid Dyad{:}Modality{:}Coupling)$$

The same model was fitted to the phase error of individual participants. Here, Subjects and Dyads were modelled as random effects.

$$Error \sim (Time + Time^2) * Modality * Coupling + (Time + Time^2 \mid Subject{:}Dyad) + (Time + Time^2 \mid Subject{:}Dyad{:}Modality{:}Coupling)$$

Statistical analyses were carried out in R (version 4.0.3); model fitting was performed with the *lme4* package (Bates, Mächler, Bolker, & Walker, 2014). All methods were carried out in accordance with relevant guidelines and regulations, based on the references provided in the respective paragraphs.

Acknowledgments

The present study was funded by *Bijzonder Onderzoeksfonds* (*BOF*) from Ghent University (Belgium).

The hardware of the recording device was built by Ivan Schepers.

Figure 2.1 was illustraded by Kevin Smink.

Professor Peter Vuust and Dr. Ole Adrian Heggli from the Center for Music in the Brain (Aarhus University, Denmark) provided precious advice and feedback on the experimental design.

Author Contributions

M.R. conceptualized, designed and implemented the study. P.J.M. supervised and contributed to study design and data processing. M.L. supervised and contributed to study design and statistical modelling. M.R. collected, processed and analyzed the data. All authors wrote the main manuscript text.

Competing interests

The authors declare no competing interests.

Data availability

All data generated during this study are included in this published article (and its Supplementary materials), together with the scripts used to analyze them.

3

Perceptual coupling in dyadic systems: continuous access to kinematics does not affect attractor dynamics

Mattia Rosso [A,B], Bart Moens [A], Canan Gener [A], Pieter-Jan Maes [A], Marc Leman[A]

[A.] IPEM - Institute for Systematic Musicology; Ghent University; Ghent, Flanders, 9000; Belgium.

[B.] PSITEC - Psychologie: Interactions, Temps, Emotions, Cognition - ULR 4072; University of Lille; Lille, Hauts-de-France, 59650; France.

Introduction

In order to move and interact with their environments, humans have to process and integrate sensory inputs from different streams of information (Spence, 2011). When the source is represented by another human, the information encoded in their movement has the potential to guide mutual adaptation of behaviors and establish minimal forms of interaction (Phillips-Silver et al., 2010). It is self-evident that the processes of accessing and decoding information from perceived rhythmic behaviors differ across sensory modalities. When we listen to someone walking, clapping or drumming, we typically hear the footstep, the clap or the hit as discrete perceptual outcomes. In the auditory domain, this class of movements tend to be characterized by naturally mute phases (A. O. Effenberg, 2005). We are 'blind' to the moment-by-moment course of their execution… until we look at how it evolves over time. Once visual contact is established, visual coupling provides individuals with continuous access to kinematic and geometric properties of an action (R. C. Schmidt et al., 1990), defined in more general terms as its informational observables (Kugler & Turvey, 2015).

Information is foundational to feed the action-perception loops underlying dyadic interactions (Tognoli et al., 2020), from the most basic level of spontaneous entrainment to more complex forms of planned joint action (Knoblich & Sebanz, 2008). Within any biological system across different scales, informational coupling is the minimal condition for individual components to behave as coordinated units. In the case of a dyadic system constituted by two coupled individuals having independent nervous systems, coupling is necessarily based on perception via one or multiple sensory channels, which sets the stage for the emergence of a dyadic control structure, formed over the informational structure of the system (Kugler & Turvey, 2015; R. C. Schmidt et al., 1990). In other words, perception enables the coordination between two individuals, and the dynamic configuration of emerging behavioral patterns depends on the specifics of the perceptual information available in the interaction. Understanding how humans access such information, and how they guide their actions accordingly, is paramount for an ecological perspective on action and perception (Gibson, 2014).

In a previous study, we reported radically different attractor dynamics in dyads engaged in a joint finger-tapping task, depending on whether participants could see (visual coupling) or hear (auditory coupling) the actions of their partner (Rosso, Maes, et al., 2021). An attractor is the dynamical entity underlying entrainment, a basic control structure (R. C. Schmidt et al., 1990) which organizes the system's behavior in coordinated states. In this and other works (Rosso, Heggli, Maes, Vuust, & Leman, 2022; Rosso, Maes,

et al., 2021), the *drifting metronomes* paradigm for dyadic entrainment was used to drive the dyad through a space of coordinative states and detect the presence of attractor points based on the relative-phase between the partners' finger-taps (from 0 to 2π radians). Despite the fact that dyadic entrainment occurred spontaneously in both sensory modalities, only in the case of visual coupling did we observe emergent coordination patterns consistently shaped by an attractor landscape (Tuller & Kelso, 1989; Yamanishi et al., 1980). More precisely, over the space explored by the drifting metronomes, the in-phase point (0,2π) stabilized the interaction in recurrent synchronized states, promoting a cooperation process in the dyadic system. The anti-phase point (π) facilitated instead decoupled independent trajectories, promoting a competition process. In the case of auditory coupling, however, the cooperation process was dominant but spread over the whole attractor landscape, with no significant influence of clear attractor points, and higher inter-dyad variability in terms of a global measure of entrainment.

Although our findings led us to conclude that attractor dynamics depend on the sensory modality mediating the coupling, we put forward the idea that the differences would attenuate or disappear if we could induce cross-modal correspondence between streams of information across visual and auditory couplings (Gallace & Spence, 2006; Marks, 2004; Parise & Spence, 2009; Spence, 2011), therefore balancing the access to movement kinematics. The observation that in everyday environments visual rhythms tend to be continuous and auditory rhythms tend to be discrete (Varlet, Marin, Issartel, Schmidt, & Bardy, 2012) biased the way we implemented the sonification of participants' finger-taps in our previous study, so that they were presented to the partner as discrete sounds at the moment of impact (Rosso, Maes, et al., 2021). Conversely, visual coupling was naturally continuous, and therefore not directly comparable to the auditory counterpart in terms of stimulus property. This limitation made it impossible to disentangle the contribution the physical property of the stimulus (continuous/discrete) from the perceptual system accessing such information (visual/auditory). We argue that these constitute two distinct dimensions of informational coupling, which require a balanced factorial design to be independently manipulated (see *Figure 3.1*).

While under ecological conditions the visual system exhibits preferential access to the observables of an action, continuously available in the optical structures it generates (R. C. Schmidt et al., 1990), sonification can be used as a means to convey kinematic information in the auditory domain (Agres et al., 2021; Bevilacqua et al., 2016; Alfred O. Effenberg, Fehse, Schmitz, Krueger, & Mechling, 2016). Auditorily coupling can be therefore artificially augmented by expanding motion acoustics to the naturally mute phases of movement (A. O. Effenberg, 2005). For finger-tapping, this would imply filling the mute inter-tap intervals by mapping movement parameters such as velocity,

acceleration, or position, onto sonic variables such as pitch, amplitude, or spatial location. Auditory presentation of continuous rhythmic cues has been shown to affect movement control in sensorimotor synchronization tasks, better guiding online trajectory, enhancing control between endpoints, and improving consistency of motor timing. This same research suggests that sensorimotor synchronization is thought to be subserved by different processes, depending on whether the synchronizer has to internally self-time the intervals or can rely on extrinsic information (Rodger & Craig, 2011). When directly comparing coordination with auditory and visual environmental rhythms, differences tend to vanish when they are presented as continuous (Varlet, Marin, Issartel, et al., 2012).

Translating these principles to the interpersonal domain, we expect to replicate for auditory coupling the same coordination patterns observed under visual coupling. Provided a 1:1 mapping of visuo-kinematic information onto a continuous sound (Alfred O. Effenberg & Schmitz, 2018), the actions of the other can be anchored to patterns of dynamic change conveyed by the sonification, such that the inter-tap intervals do not have to be explicitly computed (Maes, Giacofci, & Leman, 2015). Conversely, in a condition of visual coupling, we would expect that occluding the continuous kinematic trajectory of the finger-tapping would result in a breakdown of the coordination pattern, given that participants would lose the advantage characterizing visual modality (R. C. Schmidt et al., 1990). We refer to this scenario, wherein the patterns are explained by the access to the physical properties of motion, as the '*kinematic hypothesis*'.

Alternatively, we may observe a scenario wherein these are explained by sensory modality alone. Vision and audition evolved ontogenetically and phylogenetically with their specifics strategies for sampling environmental rhythms, and developed specific interface with motor timing system of the perceiver (Comstock et al., 2018). Being these two the main perceptual systems mediating informational coupling, we may expect that they are optimized for subserving different aspects of dyadic interactions. In this case, these differences would be reflected in different coordination dynamics across modalities, regardless of the continuous/discrete nature of the information available to the perceiver. We refer to this alternative scenario as the '*modality-dependent hypothesis*'.

The aim of the present work is to answer the question of whether interpersonal coordination relies on a supramodal processing of kinematic information ('kinematic hypothesis') or whether it depends on processing specific to the sensory modality and its unique interface with the motor system ('modality-dependent hypothesis'). To differentiate between these two scenarios, we carefully disentangled the contribution of the specific perceptual system processing the information from the physical properties of the information itself. *Figure 3.1* provides a visual representation of the experimental

design and the results we would expect from the two hypotheses. Whilst these are presented as the main expected outcomes, the design leaves room for intermediate scenarios and interactions between the two dimensions.

Results

As described in (Rosso et al., 2022; Rosso, Maes, et al., 2021), we conducted joint recurrence quantification analysis (JRQA) (Marwan et al., 2007) to compute a relational measure quantifying the degree of temporal coordination between the partners throughout the joint finger-tapping task (Marsh et al., 2009). The resulting recurrence score yielded a time-varying measure of coupling strength, quantifying instances of coordinated behavior (Marwan et al., 2007). This was used as response variable in our statistical model (see '*Methods – Statistical model*'). The score's evolution over the drifting metronomes' cycle, which depicted an attractor landscape (Tuller & Kelso, 1989; Yamanishi et al., 1980), was modeled using growth curve analysis (Mirman, 2017). For each dyad, curves were calculated as the average of 10 metronomes' cycles from 0 to 2π in every experimental condition and modeled with 2nd order orthogonal polynomials of Time. The 2-level factors, Modality (Visual, Auditory) and Kinematics (Continuous, Discrete), were modeled as fixed effects. Not having uncoupled conditions in the experimental design, effects were assessed by directly contrasting coupled conditions across levels of Modality and Kinematics. Auditory and Discrete were chosen as baseline levels for statistical comparisons, respectively.

Against our prediction, the findings provide robust evidence in favor of the '*modality-dependent hypothesis*', decisively refuting the '*kinematic hypothesis*'. This is confirmed by the significant interactions between Modality and the linear (Estimate = -1049.894, SE = 339.775, p = 0.002) and quadratic (Estimate = 1230.757, SE = 470.396, p = 0.008) terms of Time. Taken together, the effects indicate that visually coupled conditions resulted in the same asymmetrical parabolic modulation reported in (Rosso et al., 2022; Rosso, Maes, et al., 2021). The linear term of the model captures the asymmetry of the parabola, whereas the quadratic term captures the depth of the modulation. This effect was irrespective of levels of Kinematics. Furthermore, we reported a main effect of Modality (Estimate = -532.994, SE = 207.089, p = 0.010), indicating that auditory coupling resulted in higher recurrence overall, in spite of the absence of modulation over the attractor landscape. This effect holds irrespective of the levels of Kinematics. Crucially, the absence of any other effect indicates that discrete or continuous access to the partner's movements do

not differentiate across the observed attractor dynamics, nor across overall levels of recurrence.

Figure 3.2 clearly shows that the grand-average curves of the recurrence score tend to overlap within the same Modality, and are not differentiated based on Kinematics. Whereas in both Auditory conditions they fluctuate around a higher offset, in Visual conditions they form a valley with global minimum after the anti-phase midpoint. Although visual inspection may suggest that the modulation is less deep in condition of visual occlusion, the 3-way interaction between Modality, Kinematics and either polynomial terms are not statistically significant (linear: Estimate = 238.964, SE = 480.515, p = 0.619; quadratic: Estimate = 615.946, SE = 665.241, p = 0.354).

Table 3.1 presents the parameter estimates from the fixed effects model, their standard errors, and the associated p-values, with reference to the recurrence score. In this analytical framework, parameter estimates offer a measure of effect size with straightforward interpretation for linear and non-linear changes over time. By examining the interaction effect of Modality and Kinematics on the polynomial terms of Time, we could quantify the specific impact of access to movement kinematics on the evolution of the recurrence score across sensory modalities.

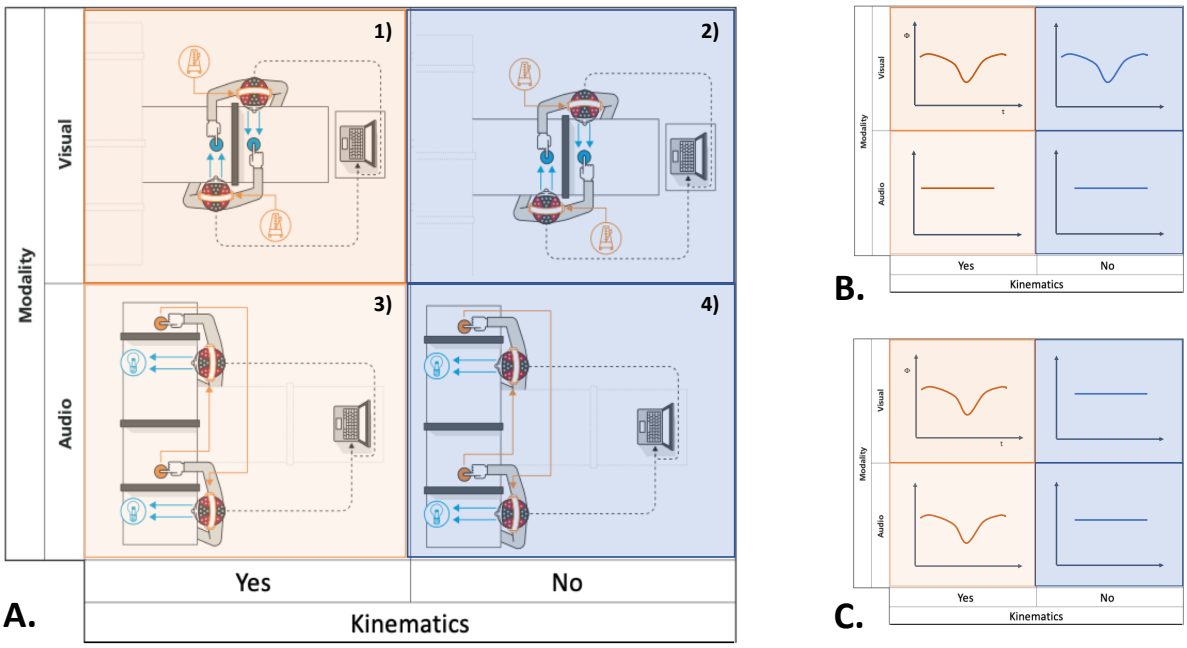


Figure 3. 1. Experimental design and hypotheses. A) Experimental design. The study was designed in a Modality (Visual, Auditory) x Kinematics (Continuous, Discrete) factorial structure. Each participant was equipped with a pad for recording their finger-tapping responses, earbuds for presentation of auditory stimuli, and a glove with an infrared reflective marker on the fingertip to continuously track the movement. A modular wooden screen with adjustable sliding windows was employed to manipulate visual coupling across conditions, allowing for various configurations (the figure serves as a conceptual representation of the manipulations and does not reflect 1:1 the physical placement of the screen and the participants). In all conditions, participants were always exposed to the partner's movements in one modality and paced by a metronome in the other modality. They were instructed to neglect the partner and focus on following their own metronome at all times. The design resulted in the following experimental conditions. **1.** ***Visually Coupled – Continuous***: participants tapped along with an auditory metronome, while looking at the partner's hand tapping. The hand of the partner was fully visible, while the view of their own hand was hidden by a component of the screen. **2.** ***Visually Coupled – Discrete***: participants tapped along with an auditory metronome, while looking at the partner's hand tapping. The hand of the partner was hidden by a sliding window, adjusted so that only the fingertip was visible at the time of impact. The view of their own hand was hidden by a component of the screen. **3.** ***Auditorily Coupled – Continuous***: participants tapped along with a flickering LED, while hearing the sonification of the partner's tapping. The sonification consisted of a continuous tone, modulated in frequency by the distance between the fingertip and the pad. A discrete 'tick' sound was triggered at the time of impact. The view of their own hand was hidden by a component of the screen. **4.** ***Auditorily Coupled – Discrete***: participants tapped along with a flickering LED, while hearing the sonification of the partner's tapping. The sonification consisted of a discrete 'tick' sound triggered at the time of impact. The view of their own hand was hidden by a component of the screen. ***B) Kinematic Hypothesis.*** This figure illustrates the expected outcomes of the joint finger-tapping task based on the 'kinematic hypothesis', wherein the coordination patterns are explained by the access to physical properties of motion (Continuous/Discrete) rather than the sensory modality (Visual/Auditory) itself. The depicted curves are purely illustrative, representing the dyad's coupling strength as a function of the attractor landscape, drawing on results from (Rosso, Maes, et al., 2021). According to this hypothesis, auditory coupling with continuous rhythmic cues would yield coordination patterns similar to those observed under continuous visual coupling, exhibiting a modulation of the interaction throughout the drifting metronomes' cycle. Conversely, visual coupling with an occluded kinematic trajectory would result in a

breakdown of the observed coordination patterns. ***C) Modality-dependent Hypothesis.*** This figure illustrates the results predicted by the complementary 'modality-dependent' hypothesis, in which coordination patterns are determined by the sensory modality (Visual/Auditory) rather than the physical properties of motion (Continuous/Discrete). Under this hypothesis, coordination patterns would differ across modalities, regardless of the continuous access to the kinematics of the partner's movements. Both figures B and C illustrate ideal scenarios entirely explained by main effects. Intermediate scenarios may emerge from interaction effects across the factors.

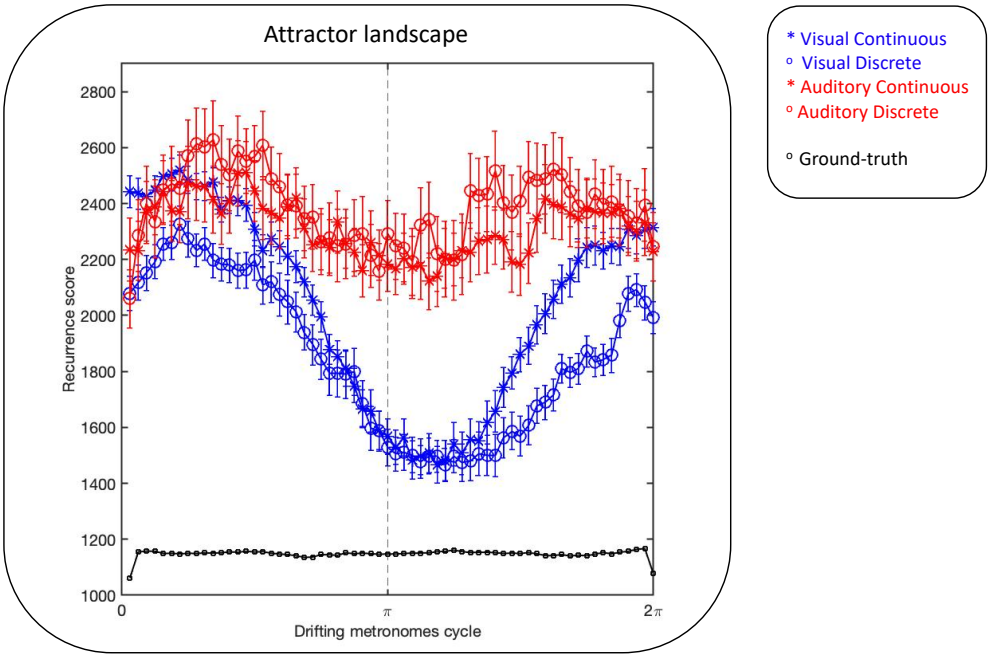


Figure 3. 2. Attractor landscape. The time series depicted in the figure represent the evolution of the recurrence score as a function of the drifting metronomes' cycle across experimental conditions. The grand average was computed over the entire sample of dyads (N=19), with each dyad's time series calculated as the average of 10 consecutive cycles. The vertical bars represent the standard error of the mean (SEM) for each phase bin. The measure is interpreted as a proxy for coupling strength between partners, and its variation over the drifting metronomes' cycle aims to provide an overview of the attractor landscape underlying the interaction. For illustration and interpretation purposes, the black line in the plot shows the same analysis performed on the two metronomes' time series. This serves as a ground truth in the context of the paradigm, providing the reference recurrence score expected by a deterministic, decoupled system such as two linearly dephasing metronomes. A horizontal line lingering at the global minimum represents the expected pattern for two partners perfectly synchronizing with their assigned metronome without influencing each other. Unlike previous studies (Rosso et al., 2022; Rosso, Maes, et al., 2021), the current experimental design does not include uncoupled conditions. Therefore, having the metronomes' collective behavior as a ground truth is essential for interpretation purposes. The black line exhibits a negligible artifact at the extremes, due to the omission of the first and the last onsets in the metronomes' time series. This artifact does not affect participants' behavior. The blue time series represent visually coupled conditions (1 and 2), where each participant tapped while seeing their partner's hand in continuous and discrete motion, respectively. Both exhibit parabolic modulation as a function of the distance from the in-phase pattern at the extremes of the metronomes' cycle, reaching a global minimum after the anti-phase midpoint. These patterns replicate the pattern previously observed in continuous visual coupling conditions, notably extending it to discrete visual coupling. While the latter suggests weaker coupling and less parabolic modulation, the effect was not statistically significant. The red time series represent auditorily coupled conditions (3 and 4), where each participant tapped while hearing the continuous and discrete sonification of

their partner's finger motion, respectively. Neither exhibited significant modulation as a function of the metronomes' cycles, but both were, on average, significantly higher than their visual counterparts. The finding replicates what was previously observed in discrete auditory coupling conditions and extends it to the case of continuous sonification based on FM. Overall, the results show that auditory coupling resulted in a higher degree of recurrence score and thus stronger coupling in the dyads. Notably, visual coupling resulted in the modulation of the recurrence score as a function of the attractor landscape. We conclude that access to continuous kinematic information, as implemented in the present experimental setup, did not significantly influence attractor dynamics, strongly supporting their dependency on the sensory modality.

	Recurrence score (N = 19)		
Predictors	*Estimate*	*SE*	*p*
(Intercept)	2387.152	168.387	0.000
Time	-155.326	265.929	0.559
$Time^2$	217.347	398.003	0.585
Modality	-532.994*	207.089	0.010
Kinematics	-75.791	207.089	0.714
Time:Modality	-1049.894**	339.775	0.002
$Time^2$:Modality	1230.757**	470.396	0.008
Time:Kinematics	-63.972	339.775	0.851
$Time^2$:Kinematics	138.860	470.396	0.767
Modality: Kinematics	257.607	292.868	0.379
Time: Modality: Kinematics	238.964	480.515	0.619
$Time^2$: Modality: Kinematics	615.946	665.241	0.354
	** p < 0.05*	*** p < 0.01*	**** p < 0.001*

Table 3. 1. Recurrence score. Orthogonal polynomials model summary.

Discussion

In the present work, we investigated the influence of kinematic information exchange between individuals on the dynamics of interpersonal synchronization, when processed in visual and auditory modalities. Our aim was to explain previous findings demonstrating that attractor dynamics depend on the sensory modality mediating informational coupling (Rosso, Maes, et al., 2021). We hypothesized that the differences could be explained by the preferential access of visual modality to continuous kinematic information conveyed by the partner's movement. Specifically, we predicted that differences in coordination patterns across modalities would cancel out, provided a reliable cross-modal match in the information available in the dyads. In order to test this hypothesis, we balanced the access to kinematics across modalities by augmenting information in the auditory domain via sonification and reducing it in the visual domain via occlusion. In doing so, we disentangled the contribution of the specific perceptual system processing information from the physical properties of the information itself.

Contrary to our hypothesis, the results suggested a genuine modality-dependency of attractor dynamics. Specifically, we found that the attractor landscape remained invariant to the kinematic properties of the partner's movement. As shown in *Figure 3.2*, the same pattern reported in conditions of visual coupling in Rosso et al., 2022; Rosso, Maes, et al., (2021) persisted even when occlusion impeded continuous access to kinematics. Conversely, the continuous sonification of movement did not modulate the interaction over time, resulting in high recurrence scores throughout the entire drifting metronomes' cycle (Rosso, Maes, et al., 2021). In summary, visually mediated interactions tended to result in periodic transitions between cooperation and competition processes modulated by opposing attractors (Marsh et al., 2009), whereas auditorily mediated interactions tended to result in constant cooperation throughout the task. We initially predicted that continuous sonification would enable participants to discern when their partner's effector was sufficiently distant from in-phase locking, allowing them to exploit the anti-phase region to decouple and pursue independent trajectories. Notably, this advantage was not exploited in the auditory modality.

We must admit, it could be argued that the adopted sonification strategy was simply not effective in conveying a continuous representation of the fingertip position with respect to the target, and consequently encoding its kinematics over time. Whilst this is certainly a possibility, our choice of frequency modulation (FM) was carefully considered and well-founded. From the perspective of the listener, the connection with low-level physical features is essential to effectively decode movement features from sound (Bresin, Mancini, Elblaus, & Frid, 2020). FM was preferred as sonification strategy over amplitude

modulation (AM) because it was proven effective in influencing auditory-motor coordination (Varlet, Marin, Issartel, et al., 2012). Whilst there seems to be no or very little difference across AM and FM in terms of stabilizing auditory-motor coordination (Rodger & Craig, 2011), FM resulted in better precision and rapidity when adopted in a 1-D guidance task (Parseihian, Gondre, Aramaki, Ystad, & Kronland-Martinet, 2016), and was proven to effectively mediate kinematic information to a perceiver when explicitly used to substitute visual information (Alfred O. Effenberg & Schmitz, 2018). Finally, pitch induces the best cross-modal correspondence with the visual dimensions of position and direction of movement along the vertical axis (Evans & Treisman, 2010; Maeda, Kanai, & Shimojo, 2004; Spence, 2011), which was the critical spatial dimension in our finger-tapping task. The mapping between the distance from the pad and the pitch was inverted, so that an ascending pitch was perceived as the fingertip approached the target. In continuously frequency modulated sounds, the detection of peaks is systematically better as compared to the troughs (de Cheveigné, 2000; Demany & Clément, 1995, 1997; Demany & McAnally, 1994), which makes an ascending modulation more encoding motion leading to a discrete event. This choice was supported by extensive piloting, and is in line with the natural correspondence with Doppler effect, ecologically experienced as sound sources approach a target. We do not exclude that results might have changed if a more effective modulation had been implemented. We address future research to investigate the impact of different sonification strategies on interpersonal coordination dynamics.

While augmenting auditory coupling through sonification offers a wide range of possibilities, there are limited approaches to visual occlusion, with the primary outcome being the obstruction of access to kinematics. Previous research has demonstrated that visuomotor synchronization significantly improves when synchronizing with moving periodic visual stimuli (Hove et al., 2010) by engaging error correction mechanisms (Hove & Keller, 2010; Hove et al., 2010). Introducing a variable velocity profile to the stimulus reduces variability in SMS tapping (Hove, Iversen, et al., 2013; Iversen et al., 2015), up to matching auditory SMS at moderate tempi (Gan et al., 2015). Given this evidence, we found it particularly surprising that visually occluding motion did not significantly disrupt the attractor pattern, nor resulted in overall weaker coupling among partners as compared to the continuous condition. Considering the invariance between continuous and discrete conditions, the distinctions may stem from the inherent characteristics of the two perceptual systems and their interactions with timing and motor mechanisms.

Beyond the interpersonal domain, the primary source of divergence in synchronization performance is typically ascribed to the distinct specializations of visual and auditory modalities, which are associated with spatial and temporal processing, respectively (Comstock et al., 2018). The auditory system's advantage in timing is thought to stem from

its stronger coupling to the motor system, which is evident in greater activation in motor structures such as the SMA and premotor cortex (Jäncke, Loose, Lutz, Specht, & Shah, 2000), as well as increased activation in the putamen (Grahn, Henry, & McAuley, 2011) independent of stimulus motion (Hove, Fairhurst, et al., 2013). Furthermore, auditory rhythms provide a selective benefit in pacing motor behavior when compensating for basal ganglia impairment in Parkinson's Disease (Ashoori, Eagleman, & Jankovic, 2015; de Dreu, van der Wilk, Poppe, Kwakkel, & van Wegen, 2012). The auditory dominance for timing is suggested to be rooted in the anatomo-functional connectivity rather than selective temporal dynamics of neural oscillations, which may not be strictly modality-specific. While activity in the beta band has been traditionally linked to timing prediction in the auditory domain (T. Fujioka, Trainor, Large, & Ross, 2009; Takako Fujioka, Ross, & Trainor, 2015; Takako Fujioka, Trainor, Large, & Ross, 2012; Iversen, Repp, & Patel, 2009), recent research has also associated beta activity with timing predictions within the visual system in response to visual rhythms (Comstock et al., 2018), as well as in cortico-kinematic coupling during real and imagined SMS (Nijhuis, Keller, Nozaradan, & Varlet, 2021). In a dyadic setting, we observed modulation of beta power as a function of the partner's movement cycles during both visual and auditory couplings (Rosso et al., 2022), extending the notion of supramodality for this neural dynamic to the interpersonal domain.

As for the visual system, despite the spatial dominance ascribed to it (Comstock et al., 2018), visual cortex exhibits activation specific to interval timing (Shuler, 2016), duration perception (B. Zhou, Yang, Mao, & Han, 2014) and prediction of rhythmic onsets (Comstock & Balasubramaniam, 2018). Despite their timing processing capabilities, these structures face additional computational demand due to the complexity of spatial processing, which is inherent to the visual scene regardless of the continuity of the stimuli. Coupled with a weaker connection to the motor system, visual targets are less likely to elicit error correction mechanisms and drive adaptive behavior (Comstock et al., 2018). Considering these differences, when visual and auditory sequences directly compete in situations where stimuli are presented out of phase during instructed synchronization with flickering lights, the auditory sequence takes precedence over the visual one, rather than the other way around (Bruno H. Repp & Penel, 2002, 2004). Such multimodal competition is directly reflected in our dyadic paradigm, where whilst sounds produced by the partner entrain individual behaviors at all times, the weaker visual coupling loses traction as the drifting metronomes drive partners away from the in-phase attractor. The conclusion we take is that the cooperation process driving dyadic entrainment, when mediated by visual coupling, strongly relies on attractor dynamics to compensate for the timing disadvantage in this modality.

How do our findings align with major theories of interpersonal synchronization (Palmer & Demos, 2022)? Predictive accounts propose that rhythmic behavior is driven by the continuous updating of internal predictive models of the other, facilitating coupled individuals to converge towards a synchronized state, with the aim of minimizing energy via error correction (Koban et al., 2019). Within this framework, our 'kinematic hypothesis' anticipated that continuous and discrete access to informational observables from a partner's movement would lead to divergent updates of predictive models. In the first scenario, the model can be updated continuously, matching the sampling rate of the specific perceptual system, thereby providing a fine-grained representation of the movement cycle. In contrast, the latter approach bases the estimation of the next tap on prior knowledge of the tapping rate, resulting in irregular, event-based updates depending on its actual occurrence. Contrary to a Bayesian account premised on supramodal updating (Fait, Pighin, Passerini, Pavani, & Tentori, 2023), our findings suggest that spatial and temporal information are differentially weighted according to the sensory input channel. For example, the 'ventriloquist effect' demonstrates the dominance of visual information in spatial processing (Alais & Burr, 2004; Radeau & Bertelson, 1974; Vroomen, Bertelson, & de Gelder, 2001), while the 'flash-sound illusion' highlights auditory supremacy in processing temporal structure (Bertelson & Aschersleben, 2003; Burr, Banks, & Morrone, 2009; Morein-Zamir, Soto-Faraco, & Kingstone, 2003; Shams, Kamitani, & Shimojo, 2000). Continuous availability of informational observables might be more heavily weighted when processed visually, as they convey information on the spatial positioning of the partner's effector. The spatial processing of continuous sonification might be minimal, and hence, have a negligible impact on behavior.

From a different perspective, within a dynamical systems framework (Demos et al., 2019; Ole Adrian Heggli et al., 2019; Stepp & Turvey, 2010), our 'kinematic hypothesis' predicted that continuous information would override the action of internal time-keepers (Richard B. Ivry & Richardson, 2002), driving behavior through simple sensorimotor interactions across modalities (Maes et al., 2015). However, the perception of oscillating objects (such as the partner's hand) varies across sensory pathways. The spatial layout of these objects is mapped onto the retina and preserved throughout the visual pathway (Luo & Flanagan, 2007; Triplett et al., 2009), whereas the receptor layout in the inner ear encodes the position of the target by combining localization and spectral cues (Middlebrooks & Green, 1991; Salminen, Tiitinen, & May, 2012). These differences reasonably lead to varied information transfer from perceptual to motor oscillators (Ole Adrian Heggli et al., 2019), suggesting that the effector is driven by unique dynamics depending on the sensory input. Consequently, our findings imply that the intrinsic dynamics of sensory perception play a pivotal role in interpersonal synchronization.

In conclusion, our study enriches the existing body of literature on interpersonal coordination by emphasizing the nature and accessibility of information, comparing the effects of visual and auditory couplings on interpersonal coordination. A thorough understanding of these facets is critical for designing and optimizing motor training and rehabilitation protocols that involve interaction between an individual and a trainer or another partner. While sonification has proven its worth for bio-feedback applications centered on the individual, both in sports (Alfred O. Effenberg et al., 2016; Lorenzoni et al., 2018; Lorenzoni, Staley, et al., 2019; Lorenzoni, Van den Berghe, et al., 2019; Maes, Lorenzoni, & Six, 2019; Van den Berghe et al., 2021) and rehabilitation (Bevilacqua et al., 2018; Maes et al., 2016; Moumdjian, Moens, Maes, Van Nieuwenhoven, et al., 2019; Van Kerrebroeck & Maes, 2021), its potential as a mediator and enhancer of dyadic coupling remains relatively unexplored. We urge further investigation into this area to broaden our understanding of the various dimensions of auditory coupling. It is essential to grasp its potential and limitations in supporting motor coordination, and to evaluate its validity for complementing visual input during dyadic interactions.

Methods

Participants. A total of 40 right-handed participants took part in the study, including 20 females and 20 males. The mean age was 31.2 years, with a standard deviation of 6.8 years. Participants were initially divided based on gender and subsequently randomly paired into 20 dyads. Upon recruitment, all participants reported not being professional musicians and having no history of neurological, major medical, or psychiatric disorders. With the exception of one dyad, participants indicated they were not familiat with their assigned partner. The experiment received approval from the Ethics Committee of Ghent University (Faculty of Arts and Philosophy), and informed written consent was obtained from each participant before commencing the experiment. Participants were compensated with a 20€ coupon for their time.

Experimental Apparatus*.* Participants were seated across from each other at the same table, on chairs equipped with armrests. By resting their elbows on the armrests, participants were able to eliminate any potential tactile or proprioceptive coupling between them due to table vibrations. Each participant was assigned a pad and instructed to tap on it using their right index finger in synchronization with a metronome. Depending on the experimental condition, the metronome provided either an auditory cue or a flickering light. Two wooden screens were placed perpendicularly on the table to manipulate visual coupling between the partners. The frontal plane was equipped with

adjustable sliding windows to regulate visual access to the partner's hand movements based on height and position of the participant. Each participant had to wear a glove with an infrared reflective marker on the finger-tip, used to track the finger's position during the task. A Qualisys Motion Capture system, consisting of 8 Miqus M3 cameras, 1 Miqus RGB camera and the 'Qualisys Track Manager 2020' software was used to track the position of the hand and tapping pads.

A Focusrite RedNet 2 Dante Audio Interface was used to deliver independent audio channels to each participant through in-ear plugs, with Ableton Live 10 serving as the main interface for stimuli presentation, sonification of finger-taps, and real-time routing. The same MIDI tracks controlled the metronomes across conditions by triggering either an audio sample or a flickering LED. A Teensy 3.2 microcontroller functioned as a serial/MIDI hub in the setup, detecting tapping onsets with < 1 ms resolution based on analog input from strain gauge pressure sensors installed within the pads. Every time a metronome onset was presented to a participant, a MIDI message was sent to the Teensy device to log the metronome's time series and control the voltage of the LEDs when necessary. Participants were monitored from a control desk placed behind a curtain using the Miqus RGB camera to ensure compliance with instructions. The setup enabled the manipulation of experimental conditions, as illustrated and described in *Figure 3.1*.

Procedure. Before participants arrived in the lab, the experimenters calibrated the motion capture system. Upon arrival in the lab, participants read all information related to the study and signed the informed consent form. Demographic data were collected with a questionnaire, and the Edinburgh inventory (Oldfield, 1971) was administered to assess participants' right-handedness. The Interpersonal Reactivity Index (IRI) (Davis & Others, 1980) was administered as a self-report measure of empathy and its subscales. Participants' preferred spontaneous tempo was calculated through 30 seconds of self-paced finger-tapping on a dedicated smartphone app *(www.beatsperminuteonline.com*). After wearing the glove for tracking the finger's position, participants wore the earbuds. Volume levels were adjusted for each participant before the experiment, and pink noise was regulated to suppress any environmental sound. After instructing the participants on how to perform finger-taps in synchronization with the metronome, they underwent a 30-seconds familiarization session. Subsequently, participants were asked to place their finger on the pad and then hold it at the highest point of their comfortable range of motion, while the experimenter recorded the positions on the software to normalize the motion range for continuous sonification. In visually coupled conditions, the lights of the room were lit at the minimum level and two diffused directional lights were oriented parallel to the table, to prevent the lights coming from the ceiling to cast the shadow of the moving hand on the pad. In condition of visual occlusion, the eight of the sliding

windows on the wooden screen was adjusted based on the height and position of the participants, so that only the fingertip of the partner was visible at the moment of impact with the pad. The importance of maintaining the position and the style of tapping was stressed in the instructions. All dyads confirmed during debriefing that the manipulation was successful. In either sensory modality (ticking auditory metronome or flickering visual metronome), each participant was cued with a different tempo (1.67 Hz and 1.64 Hz), such that the relative phase between the metronomes began at 0º, increasing regularly in 5.6º increments (Rosso, Maes, et al., 2021). A full cycle took 39 seconds to complete, with 10 consecutive cycles performed in each experimental condition. Participants were clearly instructed to ignore their partner at all times, and tap along with their assigned metronome. During breaks between experimental conditions, participants provided subjective self-reports on various aspects of the task by expressing agreement on a scale from 1 ("Completely disagree") to 7 ("Completely agree") using a custom-made battery of 12 Likert items. Data collected from the questionnaires are not presented in this paper.

Motion capture and sonification. The real-time sonification of the finger-tapping motion was implemented as follows. Five reflective markers for infrared light were attached to the pad of each participant in different configurations, such that 2 rigid bodies were identified in the motion capture recording. One extra marker was placed in the center of the pads for the creation of rigid bodies, defined as virtual marker, and physically removed to define the origin of the XYZ space. The directions of the axes were relative to the participants' bodies (X pointing outwards, Y pointing to the front, Z pointing to the upside). This way, two XY planes were defined as lying on the surface of each pad. Each participant performed the experiment wearing a glove equipped with a marker on the back of the middle phalanx of the index finger. During finger-tapping, we sonified the Euclidean distance of the marker from the XY plane of the assigned pad. This parameter was mapped onto the frequency of a sinusoidal waveform in Ableton, resulting in a continuous frequency modulation (FM) of the audio signal: upward movement away from the pad decreased the pitch, while downward movement towards the pad increased it. This sonification was implemented exclusively in Condition 3 (Auditory Coupling, Continuous) to fill the temporal gaps between tapping onsets. The time of impact was sonified as a discrete metronome sound in both conditions 3 (Auditory Coupling, Continuous) and 4 (Auditory Coupling, Discrete).

Data analysis

For the most part, the pipeline hereby described is a translation of the methods described in (Rosso et al., 2022; Rosso, Maes, et al., 2021), applied to the current experimental design. The content of this paragraph is adapted from the original work with the authors' consent. However, the values for the embedding dimension and delay differ from the original study as they were optimized for the current dataset.

Pre-processing. Throughout the duration of 10 consecutive metronome cycles, each dyad's partners were expected to produce 650 and 640 tapping onsets, respectively, over a total period of 390 seconds. Onsets occurring less than 350 ms from the previous one were considered false positives and subsequently removed, as participants could occasionally press the pad for an extended duration or inadvertently rest their hand on it. From the entire sample, 89 false positives were eliminated, accounting for 0.086% of all data points. The refined time series were then interpolated using a sine function at a 1 kHz sampling rate, yielding an estimate of the oscillators' positions within the cycle at a temporal resolution of 1 ms. The tap prior to the first metronome onset and the final tap following the last metronome onset were incorporated into the interpolation. Subsequently, data points outside the metronomes' time series boundaries were discarded. This procedure ensured equally sized time series without data loss, a prerequisite for the implementation of joint recurrence quantification analysis (JRQA; see the following paragraph). The interpolation choice is conceptually supported by the modeling of coupled oscillator systems in the context of finger-tapping studies (Ole Adrian Heggli et al., 2019; Rosso et al., 2022; Rosso, Leman, & Moumdjian, 2021; Rosso, Maes, et al., 2021). Lastly, time series were down-sampled by a factor of 4 to render the computation of recurrence plots (RPs) more manageable. As demonstrated in our previous report, JRQA results are not sensitive to the choice of down-sampling factor.

Dyad #15 was excluded from the analysis, because the Edinburgh Inventory revealed that one participant was effectively left-handed. Furthermore, the pad of one participant did not work for the whole duration of Condition #1, resulting in the loss of all data collected in that condition.

Phase-space reconstruction. Following Takens' embedding theorem (Takens, 1981), we reassembled the phase space of individual finger-tapping behaviors. This involved utilizing

time-delayed copies of the input time series *uk*, while implementing an embedding dimension *m* and a time delay *τ*.

$$\vec{x}(t) = \vec{x}_i = (u_i, u_{i+\tau}, \ldots, u_{i+(m-1)\tau}), \qquad t = i\Delta t$$

where $\vec{x}(t)$ represents the vector of reconstructed states within the phase-space at time *t*. Optimal parameters for the time-delayed embedding were calculated for each participant, for the duration of each individual metronome cycle across all experimental conditions. The resulting average value of the parameters was applied to all individual instances. The reason is that to compare recurrence rates across conditions at a group level, the embedding procedure must remain consistent among participants (e.g., see (Afsar et al., 2018; Rosso, Maes, et al., 2021), for examples of parameter selections in a factorial design). Firstly, we chose the delay *τ* as the first local minimum of the mutual information index (Fraser & Swinney, 1986) in relation to the delay. This method minimized the time series self-similarity, extracting almost orthogonal components and preventing the attractor from folding onto itself (Bradley & Kantz, 2015). The mean delay value was determined to be *τ* = 7. Next, we identified the number of embedding dimensions using the false nearest neighbor method (Rulkov et al., 1995). Specifically, we incrementally unfolded the time series into higher dimensions until data points no longer overlapped spuriously, identifying an optimal mean embedding of *m* = 2. Lastly, in line with existing literature, the maximum threshold for considering two adjacent points as recurrent was set at 10% of the maximum phase-space diameter (Marwan et al., 2007).

Joint recurrence plots (JRPs). A recurrence plot, denoted as $R_{i,j}$, is a square array employed to represent and quantify the recurrence of states within a system's phase space (Eckmann, Kamphorst, Ruelle, & Others, 1995). For each point in the phase space trajectory:

$$\vec{x}_i \ (i = 1, \ldots, N; N = n - (m\text{-}1)\ \tau)$$

the proximity to another trajectory point, $\vec{x}_j$, was assessed based on a neighborhood threshold. Individual recurrence plots are computed using the following equation:

$$R_{i,j}(\varepsilon) = \Theta(\varepsilon - \left\| x_i - x_j \right\|)$$

Here, ε represents the neighborhood threshold, $\| \cdot \|$ is the Euclidean norm (indicating the distance between two vectors), and Θ is the Heaviside step function. Each phase-space yields a square matrix with 1s where the distance $\| \cdot \|$ is less than the threshold ε, and 0s for other elements.

A joint recurrence plot (JRP) was computed for each dyad by overlapping individual RPs pairwise, retaining 1s only where both plots exhibited recurrence. Each JRP was essentially the Hadamard product of the first and second participants' recurrence plots. The *CRP Toolbox* for Matlab® was used to compute JRPs (Marwan & Kurths, 2002).

By aggregating the JRPs of the 10 trials (i.e., metronome cycles) for each experimental condition, a 2-D matrix was obtained, with each entry representing the recurrence count in the corresponding cycle region across all trials. A 1-D vector recurrence score was derived by summing the row counts for each matrix column. This vector signifies a density measure of coupled behavior instances throughout the metronomes' cycle.

The recurrence score scale relies on the JRP size, which, in turn, depends on the embedding procedure. This necessitates setting consistent parameters across the entire sample. To enhance the signal-to-noise ratio and prevent oversampling, the resulting time series were reduced to 64 bins by averaging recurrence scores for equally sized, consecutive time periods. The interval size was determined by the slower metronome's increments, which provided a regular subdivision inherent to the experimental trials.

All processing steps were conducted in Matlab®. Our approach was chosen over the moving window-based JRQA version to avoid low-pass filtering effects on the time series, which could impede result interpretation. Specifically, a moving window can introduce phase distortion in the time series based on window size, making it unreliable for detecting attractor points across the landscape.

Statistical models. The recurrence score served as the response variable in a mixed-effects model, which included Coupling and Perspective as factors and Time as a continuous predictor, expressed through the indexes of the metronome steps (from 1 to 64). Due to the non-linear time course observed in coupled conditions, we employed the method of orthogonal polynomials (Mirman, 2017), incorporating linear and quadratic functions of

Time into our model. Dyads and interactions between Dyads and the factors were modeled as random effects on all polynomial terms to accommodate individual variability in synchronization abilities and individual susceptibility to coupling across experimental manipulations. The random effects structure was adopted to minimize false alarm rates without significant power loss (Barr et al., 2013). Guided by our prior study and the examination of empirical curves from the current dataset, we confined the polynomial model to the 2nd order as the most parsimonious solution. Within this analytical framework, the intercept is considered a 'zero-order' polynomial, as it displays zero changes in any direction. Significant changes in direction signify modulation by the task's temporal structure, enabling us to quantify the influence of attractor points as the dyad diverged from the horizontal trajectory while transitioning over anticipated critical regions.

The full model's formula is as follows:

$$Recurrence \sim (Time + Time^2) * Modality * Kinematics + (Time + Time^2 \mid Dyad) + (Time + Time^2 \mid Dyad{:}\,Modality{:}\,Kinematics)$$

Statistical analyses were carried out in R (version 4.0.3). The *lme4* package (Bates et al., 2014) was used for model fitting.

4

Embodied perspective-taking enhances interpersonal synchronization. A body-swap study.

Mattia Rosso [A,B], Bavo van Kerrebroeck [A,C,D], Pieter-Jan Maes [A], Marc Leman [A]

[A.] IPEM - Institute for Systematic Musicology; Ghent University; Ghent, Flanders, 9000; Belgium.

[B.] PSITEC - Psychologie: Interactions, Temps, Emotions, Cognition - ULR 4072; University of Lille; Lille, Hauts-de-France, 59650; France.

[C.] SPL - Sequence Production Lab; McGill University; Montréal, Québec, H3A 1B1; Canada.

D IDMIL – Input Devices. And Music Interaction Laboratory; McGill University; Montréal, Québec, H3A 1E3; Canada

Rosso, M., van Kerrebroeck, B., Maes, P. J., & Leman, M. (under review). Embodied perspective taking enhances interpersonal synchronization. iScience

.

Introduction

Humans exhibit a compelling tendency to synchronize rhythmic movements with one another (Crombé et al., 2022; Richardson, Marsh, & Baron, 2007; Richardson et al., 2005; R. C. Schmidt & O'Brien, 1997). As soon as two individuals exchange information via one or multiple sensory channels (R. C. Schmidt et al., 1990), such phenomenon may occur spontaneously and even against the intention to ignore the other (Rosso, Maes, et al., 2021). Visually-mediated interactions, in particular, are governed by attractor dynamics (J. A. S. Kelso, 1995; Tognoli et al., 2020) which stabilize dyadic behavior in recurrent and stable coordinative patterns, and are characterized by a dynamic balance between the pursuit of individual behavioral trajectories (*competition* process) and the attraction into coupled behavior (*cooperation* processes) (Marsh et al., 2009; Rosso, Maes, et al., 2021).

Ecological dyadic interactions take place in settings where individuals perceive each other from a face-to-face 2nd person perspective. However, embodied simulation accounts of social cognition posit that the mirroring of another person's movements is enabled by neural representations based on a bodily format (Gallese & Sinigaglia, 2011), which require a visuospatial transformation to remap the observed movement into an egocentric frame of reference (Oh, Braun, Reggia, & Gentili, 2019). Despite broad evidence for such form of embodied perspective-taking (Graf, 1994; Keehner, Guerin, Miller, Turk, & Hegarty, 2006; K. Kessler, 2000; Klaus Kessler & Thomson, 2010; May, 2004; Wraga, Shephard, Church, Inati, & Kosslyn, 2005; Zacks & Michelon, 2005), the mechanism remains overlooked in the literature on dyadic interactions (Oh et al., 2019). In the present work, we investigate the role of visual perspective in interpersonal temporal coordination, under the hypothesis that perceiving the movements of a partner from their 1st person perspective would promote their motor alignment, therefore strengthening and stabilizing their synchronization.

Perspective-taking can nowadays be induced in an embodied bottom-up fashion, by experimentally transposing the visual scenes perceived by two individuals into the partner's egocentric frame of reference (Petkova, Björnsdotter, et al., 2011). Manipulations of this kind tap into the plasticity of body schemas as represented in the central nervous system (Tsakiris, 2010, 2017), and effectively lead to experience embodiment (Botvinick & Cohen, 1998; K. Kilteni, Groten, & Slater, 2012; Slater, Perez-Marcos, Ehrsson, & Sanchez-Vives, 2008, 2009) and agency (Kalckert & Ehrsson, 2012, 2014) over effectors not belonging to one's own body, as long as they are visually perceived in a configuration which is coherent with bodily constraints (Pavani, Spence, & Driver, 2000). The same principle, originally investigated by means of the rubber hand

illusion (Botvinick & Cohen, 1998), was extended to the more radical experience of full-body (Maselli & Slater, 2013; Petkova, Björnsdotter, et al., 2011; Petkova, Khoshnevis, & Ehrsson, 2011; Serino et al., 2013; Slater et al., 2009) and out-of-body (Ehrsson, 2007; Guterstam & Ehrsson, 2012) illusions, where a person gets to experience ownership and agency over a humanoid virtual avatar in 1st person or dislocation respect to the position of the real body, respectively. Applied in a social setting, the same principles allow to induce a full body-swap between two real persons by immersing the streaming the 1st person view of one partner into the visual scene of the other (Petkova & Ehrsson, 2008).

Konban et al. (Koban et al., 2019) proposed that dyadic synchronized behavior is guided by an optimization principle, aimed at minimizing prediction errors by correcting the temporal mismatch between movements executed by one self and movements executed by the other. In terms of brain-body-environment system, motor control is guided by environmental contingencies towards a reduction of computational cost (Clark, 1999). With these principles in mind, let us take the human hand as paradigmatic effector to investigate embodiment (Botvinick & Cohen, 1998), and joint finger-tapping as paradigmatic task to investigate interpersonal synchronization (Ole A. Heggli et al., 2019; Konvalinka et al., 2010; Rosso, Maes, et al., 2021). Provided the hand of a partner can be integrated in one's own body schema when visually perceived in 1st person during joint finger-tapping (Dell'Anna et al., 2018), we expected temporal mismatches to carry more weight as compared to the ecological 2nd person perspective, which would lead the partners to engage in a stronger error-correction response via (dyadic) entrainment. This is because, in such scenario, the other's hand is perceived as an embodied effector and therefore represented by the motor system in terms of motor potentiality for actions (Della Gatta et al., 2016; Gallese & Sinigaglia, 2010). When the visual feedback expected from a motor output is altered, resulting in a mismatch with the prediction of a forward-model (S. J. Blakemore, Frith, & Wolpert, 1999; S.-J. Blakemore, 2017; Wolpert, Doya, & Kawato, 2003), humans spontaneously engage in motor adaptation to keep a consistent relationship between action and perception (Maes et al., 2015). This brings us to our central research questions. 1) Can we induce spontaneous interpersonal synchronization by visually coupling, in 1st person perspective, two individuals engaged in a joint-finger tapping task? 2) Can visual coupling in 1st person strengthen interpersonal synchronization, as compared to the ecological 2nd person perspective? 3) How do attractor dynamics compare across different visual perspectives?

In order to answer these questions, we adopted the 'drifting metronomes' paradigm for dyadic entrainment (Rosso, Maes, et al., 2021) and manipulated the visual scenes perceived by the partners across experimental conditions. The paradigm consists of a joint finger-tapping task, where two partners are instructed to entrain with incongruent

auditory rhythms, while being visually coupled with one another. Despite the explicit instruction to ignore each other, recurrent patterns of spontaneous coordinated behavior emerge between individuals according to consistent temporal dynamics (Rosso, Maes, et al., 2021). The essential feature of the paradigm is a minimal gap in the frequencies of the two metronomes, such that when they are set to start at the same time, their relative phase systematically increases with every beat from 0 to π radians and subsequently decreases from π to 0. Cyclical repetitions of such pattern allow to identify regions of maximal attraction over the whole attractor landscape (Tuller & Kelso, 1989), capturing the time-varying nature of dyadic entrainment beyond a global measure of synchronization. The same task was performed under different conditions of body-swap. As illustrated in *Figure 4.1*, during the task participants were either seeing the other's hand in 2nd person (1), their own hand in 1st person (2), the other's hand in 1st person (3) or their own hand in 2nd person (4). The subjective feeling of embodiment was measured via a visuotactile stimulation procedure based on the principles of the rubber-hand illusion. Experimental design and procedures are described in detail in the *Methods* section.

Finally, we tested the well-documented association between interpersonal synchronization and empathic traits (for a recent review, see Tzanaki, 2022), and with the self-reported sense of ownership over the other's hand. At a higher cognitive level, a bottom-up driven experience of being 'in the shoes of the other' mitigates outgroup (Maister, Slater, Sanchez-Vives, & Tsakiris, 2015) and racial biases (Peck, Seinfeld, Aglioti, & Slater, 2013; Thériault, Olson, Krol, & Raz, 2021), attenuates gender stereotype threat (Peck, Doan, Bourne, & Good, 2018; Peck, Good, & Bourne, 2020), promotes perceived self-other similarity (Paladino, Mazzurega, Pavani, & Schubert, 2010; Tajadura-Jiménez, Grehl, & Tsakiris, 2012) and even the social acceptability of a humanoid robot (Ventre-Dominey et al., 2019). Crucially, both empathy and synchronization activate embodied representations of observed actions in the brain (Gallese, 2019), while high trait empathy (Novembre, Ticini, Schütz-Bosbach, & Keller, 2012) and empathic perspective taking (Novembre et al., 2019) was shown to strengthen such representations. We therefore hypothesized that both high scores in cognitive perspective taking and subjective experience of embodiment would predict stronger entrainment with the partner, in particular when assuming their 1st person visual perspective.

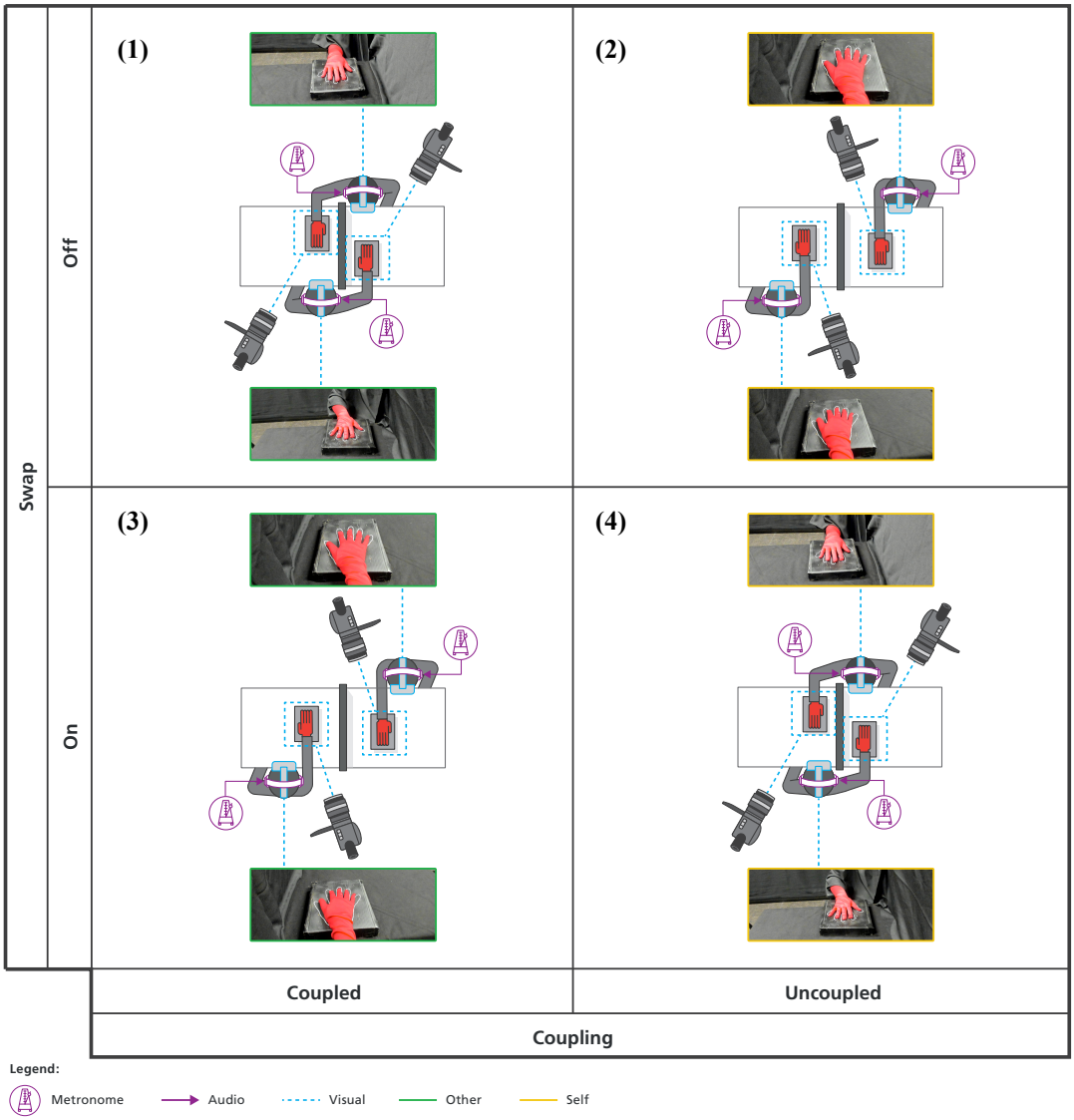


Figure 4. 1. Experimental design. The study was designed in a Perspective (2P, 1P) x Coupling (Coupled, Uncoupled) factorial structure. Each participant was equipped with a headset providing full immersion in different visual scenes across experimental conditions. Visual scenes were captured and streamed in real-time by cameras placed either in front of the partner's hand or above the participant's shoulder, as illustrated in the detail boxes of the figure. The setup allowed for the crucial manipulation of swapping the visual scenes as captured by different angles, illustrated in the schema as 'Swap On/Off'. The design resulted in the following experimental conditions. 1) *'2P Coupled'*. Participants tapped along with an auditory metronome, while looking at the partner's hand tapping. The partner's hand was video-recorded from a frontal position and no swapping was performed, so that the hand was perceived from a 2nd person perspective. Participants were explicitly asked to neglect the partner's movements and focus on following their own metronome. 2) *'1P Uncoupled' (control).* Participants tapped along with an auditory metronome, while looking at their own hand. Their own hand was video-recorded from above their shoulder and no swapping was performed, so that the hand was perceived from a 1st person perspective. 3) *'1P Coupled'*. Participants tapped along with an auditory metronome, while looking at the partner's hand tapping. Their own hand was video-recorded from above their shoulder, but swapping was performed so that the partner's hand was perceived from a 1st person perspective (as the partner would see oneself). Participants were explicitly

asked to neglect the partner's movements and focus on following their own metronome. 4) ***'2P Uncoupled' (control).*** Participants tapped along with an auditory metronome, while looking at their own hand. The partner's hand was video-recorded from a frontal position, but swapping was performed so that the own hand was perceived from a 2nd person perspective (as oneself would be seen by the partner).

Results

Taking dyadic behavior as unit of analysis, we replicated joint recurrence quantification analysis (JRQA) (Marwan et al., 2007) as presented in Rosso et al. (Rosso, Maes, et al., 2021). The approach yielded a relational measure to quantify the degree of temporal coordination between the partners throughout the task (Marsh et al., 2009). A recurrence score was calculated as time-varying measure of coupling strength, quantifying the instances of coordinated behavior (Marwan et al., 2007). Its evolution over the drifting metronomes' cycle, portraying an attractor landscape (Tuller & Kelso, 1989), was modelled by means of growth curve analysis (Mirman, 2017). For every dyad, curves were calculated as the average of 10 metronomes' cycles from 0 to 2π in each experimental condition, and modeled with 2nd order orthogonal polynomials of Time. 2-level factors Coupling (Coupled, Uncoupled) and Perspective (1P, 2P) were modelled as fixed effects. Uncoupled (control) conditions were treated as baseline for contrasting levels of the Coupling factor, 2P conditions were treated as baseline for the Perspective factor.

In line with our hypotheses, we found a significant main effect of Coupling (*Estimate* = 436.278, *SE* = 73.777, $p < 0.001$) indicating an overall increase of the recurrence score in presence of informational coupling as compared to the uncoupled control conditions, independently from the manipulation of Perspective. We also found a significant interaction effect between Coupling and the quadratic term of Time (*Estimate* = 1201.669, *SE* = 166.386, $p < 0.001$), meaning that the modulation of the attractor landscape on the response variable resulted in a significant 'valley' around the anti-phase midpoint in coupled conditions. Crucially, we found a 3-way interaction between the linear component of Time, Coupling and Perspective (*Estimate* = 478.850, *SE* = 215.621, $p = 0.026$), indicating that the linear coefficient in coupled conditions significantly differed across 1st person and 2nd person levels of Perspective. Such interaction captures the change in the asymmetry of the parabolic curves across the two Coupled conditions. Whilst in 2nd person the recurrence score lingers in the 'valley' into the second half of the drifting metronomes' cycle, in 1st person it bottoms at the anti-phase point and grows

straight towards the in-phase point. *Figure 4.2* shows the grand-average curves of the recurrence score across experimental conditions.

Table 4.1 shows the fixed effects parameter estimates and their standard errors for recurrence score, along with associated p-values. In this analysis framework, parameter estimates provide a measure of effect size of straightforward interpretation for linear and non-linear changes over time, as long as the polynomial order is not too high (Mirman, 2017). With the interaction effect of Coupling and Perspective factors on the polynomial terms, we could quantify the specific effects of visual perspective on the evolution of the recurrence score.

Moving at the individual level of analysis, we tested whether the manipulation of visual perspectives induced significant changes in the experienced sense of ownership over the perceived hand. Aligned rank transform (ART) ANOVA (Wobbrock, Findlater, Gergle, & Higgins, 2011) revealed significant main effects of Coupling (*Df residual* = 147, *F* = 104.353, $p < 0.001$) and Perspective (*Df residual* = 147, *F* = 8.983, $p < 0.01$) on the self-reported ownership ratings. The former indicates that participants were capable of telling apart their own hand from the partner's regardless of the visual perspective, whilst the latter indicates that perceiving a hand in 1st person generally resulted in a stronger sense of ownership. Crucially, the interaction effect between Coupling and Perspective (*Df residual* = 147, *F* = 5.232, $p < 0.05$) revealed that the increase in ownership relative to the 2nd person perspective was significantly stronger when participants were coupled. This means that the partner's hand, normally recognized as belonging to somebody else, is perceived as belonging to one's own to a significantly greater extent due to the manipulation of visual perspectives. The same model was fit to the ratings of sense of ownership and sense of agency as experienced during the joint finger-tapping task. In this case, we only found significant main effect of Coupling on both ownership (*Df residual* = 147, *F* = 459.467, $p < 0.001$) and agency (*Df residual* = 147, *F* = 373.005, $p < 0.001$), indicating that in conditions of active movement participants correctly attributed the hand and its actions to their selves, and did not experience illusory attribution of the partner's hand from any visual perspective. Median scores for self-reported ratings are shown in *Figure 4.3*.

Finally, the *perspective taking*, *empathic concern*, *fantasy* and *personal distress* scores subscales of empathy (Davis & Others, 1980) had no significant effect on synchronization consistency, computed as the resultant vector length R of relative phase with the assigned metronome. However, the same model revealed a significant 2-way interaction between Coupling and Perspective (*Estimate* = -0.292, *SE* = 0.141, $p < 0.05$), indicating that the negative impact of coupling on individual synchronization performance with the metronome was strengthened by perception 1st person perspective. This corroborates the results from the dyadic analyses, which showed that attraction towards the in-phase point was stronger in such condition. As demonstrated in Rosso et al. (Rosso, Maes, et al., 2021),

a reduced synchronization with the metronome in Coupled conditions as contrasted to Uncoupled conditions is a valid proxy for the attraction towards the partner, at the expenses of the compliance with the task.

The same models fitted on the subjective ratings of embodiment revealed a significant 2-way interaction effect between *empathic concern* and Coupling (*Estimate* = -0.177, *SE* = 0. 089, *p* = 0.05). This indicates that participants with higher empathic concern experienced a stronger sense of ownership over the partner's hand. We also point out a trend towards 3-way interaction with Coupling and Perspective (*Estimate* = 0.207, *SE* = 0.128, *p* = 0.11), showing that the effect tended to be stronger when the partner's hand was perceived in 1st person. Models' summaries for the Individual level of analysis are reported in *Table 4.2, 4.3 and 4.4*. In all models, Uncoupled (factor Coupling) and 2P (factor Perspective) were set as 0-levels for statistical contrasts.

	Recurrence score (N = 19)		
Predictors	*Estimate*	*SE*	*p*
(Intercept)	752.739	56.693	0.000
Time	-106.503	120.749	0.378
$Time^2$	4.933	127.602	0.969
Perspective	-0.180	73.777	0.998
Coupling	436.278 ***	73.777	< 0.001
Time:Perspective	53.653	152.467	0.725
$Time^2$:Perspective	49.110	166.386	0.768
Time:Coupling	-122.145	152.467	0.423
$Time^2$:Coupling	1201.669 ***	166.386	< 0.001
Perspective:Coupling	50.670	104.337	0.627
Time:Perspective:Coupling	478.850 *	215.621	0.026
$Time^2$:Perspective:Coupling	7.047	235.305	0. 976
	** p < 0.05*	*** p < 0.01*	**** p < 0.001*

Table 4. 1. Recurrence score. Orthogonal polynomials model summary.

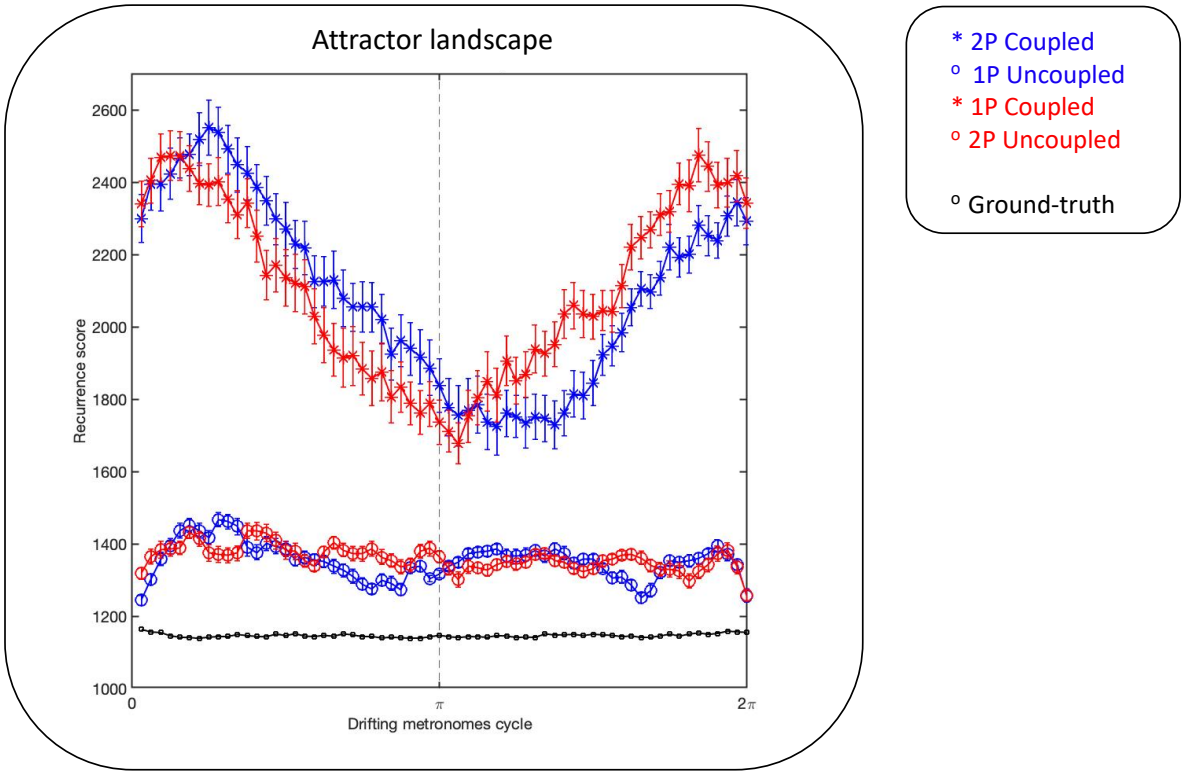


Figure 4. 2. Attractor landscape. The timeseries depicted in the figure represent the evolution of the recurrence score as a function of the drifting metronomes' cycle, across experimental conditions. The grand-average was computed over the whole sample of dyads (N=19), and for each dyad the timeseries was computed as the average of 10 consecutive cycles. For illustration and interpretation purposes, the black line in the plot shows the same analysis as performed on the two metronomes timeseries. This represents a ground-truth in the context of the paradigm, providing the reference recurrence score expected by a deterministic de-coupled system such as two linearly dephasing metronomes. A horizontal line lingering at the global minimum is the pattern expected from two partners when each of them is perfectly synchronizing with the assigned metronome, without influencing each other. The two timeseries just above the reference were computed from the uncoupled conditions (2 and 4), where each participant was tapping while seeing their own hand from 1st and 2nd person perspectives, respectively, so that no information was exchanged with the partner. As expected, the recurrence scores closely tracked the reference in these conditions, with random fluctuations around the mean and spurious recurrences due to human movement variability. Due to absence of coupling between the partners and hence their ignorance of the drifting metronomes' structure, it was feasible for them to follow the assigned metronomes. No significant difference was found between visual perspectives in uncoupled conditions. These control conditions provided a baseline for statistical contrasts, allowing us to assess the significance of eventual patterns deviating from the reference due to visual coupling. The two upper timeseries represent the coupled conditions (1 and 3), where each participant could see the hand of the partner from 2nd and 1st person perspectives, respectively. These are the critical conditions to focus on, in order to answer our main research question. When modelling empirical curves with orthogonal polynomials (Mirman, 2017), the intercepts of the fitted model were significantly greater than the uncoupled control conditions, capturing the main effect of Coupling. It is indeed clearly visible that both curves are on average above the respective controls. Focusing on the shape of the curves, it is also evident that both exhibit a parabolic curvature and a pronounced asymmetry. These two features were captured by the significant interaction effects of Coupling with the Quadratic and the Linear terms of Time, respectively.

Neither the average recurrence score nor the parabolic curvature significantly differed across levels of Perspective. However, we did find a significant interaction on the Linear term of Time, capturing a critical difference between the curves depending on visual perspective. Whilst in both coupled conditions the recurrence score reached a global minimum past the π midpoint, in 2nd person perspective it lingered on a longer horizontal trajectory into the second half of the cycle, before reaching the maximum with a steeper exponential growth. This resulted in a more prominent asymmetry which, as previously discussed (Rosso, Maes, et al., 2021), indicates hysteresis in the system (Marsh et al., 2009; Richardson, Marsh, & Baron, 2007) since the rate of change of the recurrence score is dependent on the direction of the de-phasing (i.e., from 0 to π and from π to 0). The same in-phase attractor exerted a stronger 'pull-back' on the dyad as it left the 0 point, followed by a steeper 'push-forward' as it approached the same point at the end of the cycle. The anti-phase point can be seen as a 'competition attractor' (Marsh et al., 2009), for it facilitates de-coupling among the partners and pursuing of independent trajectories. This interpretation is empirically supported by our reference timeseries (black line in the plot), showing that horizontal line at a minimum occur when a de-phasing pattern is taking place. Crucially, the competition attractor around the π point resulted to be weaker when partners were coupled in 1st person perspective, since they did not manage to keep dephasing for quite as long. The dynamic balance shifted in favor of the cooperation attractor, resulting in a steeper increase of recurrence score. From these observations, we conclude that visual coupling in 1st person promotes the convergence of the dyadic system towards phase-alignment as compared to the ecological mode of interaction in 2nd person.

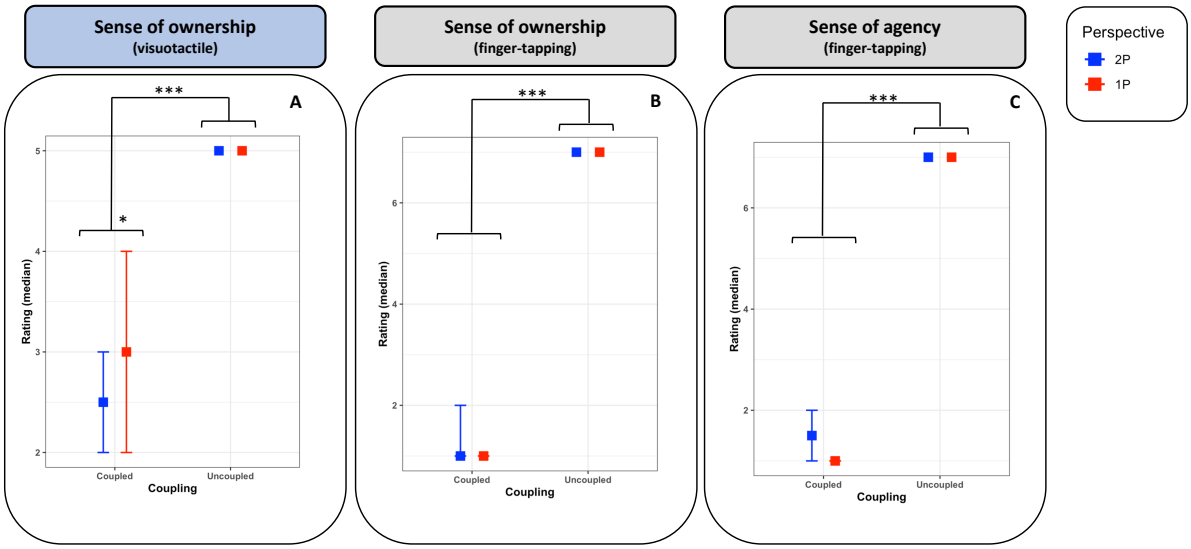


Figure 4. 3. Sense of ownership over the perceived hand. Boxes indicate median values for the subjective ratings referring to the following constructs, across experimental conditions: (A) sense of ownership during visuotactile stimulation (scale 1-5), (B) sense of ownership during finger-tapping task (scale 1-7), (C) sense of agency during finger-tapping task (scale 1-7). Error bars indicate the 95% confidence interval of the median. The fact that in Uncoupled conditions participants systematically recognized the hand as their own resulted in a ceiling effect, which did not allow to compute error bars in such conditions. For all response variables (A, B, and C), we found a main effect of Coupling, whereas the interaction effect was significant only when sense of ownership was measured via visuotactile stimulation (A). Whilst the main effect is somewhat trivial, the interaction shows that the manipulation of visual perspectives was successful in inducing a subjective experience of embodiment over the partner's hand, specifically when this was perceived from a 1st person visual perspective.

	Sense of ownership – visuotactile stimulation (N = 19)		
Predictors	*Df residual*	*F value*	*p*
Coupling	147	104.353***	< 0.001
Perspective	147	8.983**	0.003
Coupling:Perspective	147	5.232*	0.023
	** p < 0.05*	*** p < 0.01*	**** p < 0.001*

	Sense of ownership – finger-tapping task (N = 19)		
Predictors	*Df residual*	*F value*	*p*
Coupling	147	459.467***	< 0.001
Perspective	147	2.356	0.127
Coupling:Perspective	147	1.278	0.260
	** p < 0.05*	*** p < 0.01*	**** p < 0.001*

	Sense of agency – finger-tapping task (N = 19)		
Predictors	*Df residual*	*F value*	*p*
Coupling	147	373.005***	< 0.001
Perspective	147	0.063	0.802
Coupling:Perspective	147	0.021	0.886
	** p < 0.05*	*** p < 0.01*	**** p < 0.001*

Table 4. 2. Sense of ownership. ART 2-way ANOVA models summaries.

	Vector length R (N = 38)		
Predictors	*Estimate*	*SE*	*p*
(Intercept)	0.815	0.101	0.000
Empatic concern (EC)	0.003	0.007	0.618
Coupling	-0.139	0.098	0.160
Perspective	-0.038	0.098	0.694
Coupling:Perspective	-0.292*	0.140	0.041
EC:Coupling	0.002	0.007	0.697
EC:Perspective	0.005	0.007	0.486
EC:Coupling:Perspective	0.016	0.009	0.100
	** p < 0.05*	*** p < 0.01*	**** p < 0.001*

Table 4. 3. Vector length R (synchronization with metronomes). Mixed-effects model summary.

	Sense of ownership (N = 37)		
Predictors	*Estimate*	*SE*	*p*
(Intercept)	4.768	1.016	0.000
Empatic concern (EC)	-0.013	0.068	0.850
Coupling	0.643	1.333	0.6306
Perspective	-0.184	1.333	0.890
Coupling:Perspective	-2.922	1.908	0.129
EC:Coupling	-0.177	0.089	0.050
EC:Perspective	0.025	0.089	0.779
EC:Coupling:Perspective	0.207	0.128	0.110
	** p < 0.05*	*** p < 0.01*	**** p < 0.001*

Table 4. 4. Sense of ownership. Mixed-effects model summary.

Discussion

The present study investigated the role of visual perspective in spontaneous dyadic entrainment, which is considered to be the most minimal and fundamental level of rhythmic interpersonal coordination (Knoblich & Sebanz, 2008; Marsh et al., 2009; Sebanz & Knoblich, 2009). By inducing a hand-swap illusion between partners engaged in joint finger-tapping, we were able to quantify overall synchronization strength and local attractor dynamics when they could perceive each other's hand in 1st person, and compare them to an ecological mode of interaction in 2nd person. The drifting metronomes paradigm was adopted to guide the partners through a systematic exploration of their attractor landscape, to detect attractor points over the whole space of coordinative states (Rosso, Maes, et al., 2021). As analysis framework, joint recurrence quantification analysis (JRQA) (Marwan et al., 2007) yielded a relational measure to quantify the degree of temporal coordination within the dyad throughout the task (Marsh et al., 2009).

In the first place, we were able to replicate the results from our previous report (Rosso, Maes, et al., 2021). When participants were visually coupled in a 2nd person face-to-face interaction, a cooperation process dominated the interaction, resulting in recurrent states of coordinated behavior despite the active attempt of neglecting the partner's rhythm and pursue individual trajectories. Crucially, the effect was not constant over the whole drifting metronomes' cycle, but rather modulated by consistent attractor dynamics. As dyads were driven by the metronomes through the space of relative phase values, we could observe the recurrence score oscillating between global maxima and global minima in proximity of critical regions. High recurrence score indicates high degree of temporal coordination within the dyad, while low recurrence score indicates temporal independence. When the recurrence score lingers at baseline levels for a sustained period of time, it means that partners managed to ignore each other and maintain their own tempo, tracking the de-phasing pattern of the drifting metronomes. It is in the transitions over these critical regions that the dynamic balance between two opposite tendencies of the system can be observed, namely the 'pull' into temporally coordinated behavior and the 'push' towards de-coupled, independent behavior. The maxima and minima of the recurrence score were observed around the in-phase (0) and anti-phase (π) points, which operated as 'cooperation attractor' and 'competition attractor' (Richardson, Marsh, & Baron, 2007), respectively. Whilst the partners tended to move together at a collective level of coupled behavior in proximity of the cooperation attractor, it became easier for them to pursue independent de-coupled trajectories in proximity of the competition attractor.

Moving on to the present experimental design, the main effect of Coupling revealed that recurrence score was on average significantly higher in coupled conditions (global cooperation), while its 2-way interaction with the quadratic component of Time showed a significant modulation by attractor points over the course of the drifting metronomes' cycle (local dynamics). More specifically, recurrence score grew as a parabolic function of the relative phase between metronomes, finding its maximum around the in-phase point and its minimum right after the middle anti-phase point. As showed in *Figure 4.2*, when participants were visually coupled, both 1st and 2nd person perspectives scored on average above the baseline levels of the uncoupled control conditions, and exhibited the same depth of the parabolic curvature. Crucially, although the mutual assumption of 1st person perspective did not affect the global level of recurrence, it resulted in a stronger attraction towards the cooperation attractor in the second half of the cycle. The effect was captured by the significant 3-way interaction between Coupling, Perspective, and the linear term of Time, which indicates that the asymmetry in the parabolic curve significantly changes depending on the levels of perspective. Due to the manipulation, the pull towards the in-phase attractor began earlier on in the cycle, such that participants did not manage to take advantage of the π region to de-couple. In our paradigm, the relative phase between the metronomes was manipulated as control parameter (S. Strogatz et al., 1994) from 0 to π (ascending) and from π to 0 (descending) radians. In this scenario, dyads tended to slowly transition from cooperation to competition as the metronomes' relative phase diverged from 0 to π, whereas they exhibited a more abrupt transition from competition to cooperation as metronomes converged from π to 0. The asymmetry, observed in condition of 2nd person and captured by the linear term of Time in the polynomial model, was previously reported in the paradigm (Rosso, Maes, et al., 2021) as manifestation of hysteresis, namely the dependency of the dyadic system on history and directionality of the interaction (Richardson, Marsh, & Baron, 2007). We therefore interpret a steeper growth of the recurrence curves as a sign of increased hysteresis and stronger cooperation attractor when participants were coupled in 1st person perspective (Marsh et al., 2009).

When explaining our finding, we first have to rule out that any differences in coupling strength could be explained by varying amounts of information. In fact, except the fact that it was rotated by 180°, the physical features of the observed hand were held constant across conditions (see *Figure 4.1*). Given the hand orientation with respect to the observer's body was the only discriminant, we should ground our explanation on how the brain represents a bodily effector when centered on egocentric frame of reference, and why this would facilitate motor adaptation to temporal mismatches between performed and observed actions, resulting in stronger entrainment as emergent property of dyadic behavior. We propose two possible explanations for this. The first one, based on social

cognition, frames interpersonal synchronization in terms of self-other integration (Sarah Jayne Blakemore & Frith, 2003; G. Dumas, Moreau, Tognoli, & Kelso, 2020; Farrer & Frith, 2002; Ole Adrian Heggli et al., 2019; Ole Adrian Heggli, Konvalinka, Kringelbach, et al., 2021; Huberth et al., 2019; Koban et al., 2019; Novembre, Sammler, & Keller, 2016; van der Meer, Groenewold, Nolen, Pijnenborg, & Aleman, 2011; Varlet et al., 2020), where distinct but overlapping brain networks (Farrer & Frith, 2002; C. D. Frith & Frith, 1999; Hasson & Frith, 2016; Mitchell, Banaji, & Macrae, 2005) process information related to movements produced by the self and information stemming from movements observed in the other. To account for dynamic development of dyadic rhythmic interactions, it was recently proposed that self-other integration and segregation are two metastable attractor states underlying coupled and de-coupled behaviors, respectively (Ole Adrian Heggli, Konvalinka, Kringelbach, et al., 2021). Such metastability between integration/segregation states maps well onto the cooperation/competition attractors we discussed so far. According to this view, transitions occur as the brain selects whether it is more efficient to integrate perceptual information in one merged model for self and other, or hold two segregated models to attribute the perceived action to one of the two agents. The visuospatial overlap between the other's hand perceived in 1st person and one's own hand would blur the difference between self and other, promoting the merging of two separate models into one, tightening action-perception loops between the partners, and ultimately strengthening the attraction towards a coupled sate.

The second explanation is based on a purely sensorimotor account, it does not call into question the representation of the other, and is based on the fact that our manipulation set the conditions for the brain to represent the perceived effector as actually belonging to the bodily self (Botvinick & Cohen, 1998; Della Gatta et al., 2016; Kalckert & Ehrsson, 2014; Slater et al., 2008, 2009; Tsakiris, 2010, 2017; Tsakiris & Haggard, 2005). The mapping between actions and their sensory consequences is learned throughout a lifetime of sensorimotor contingencies and action-perception dependencies (Shadmehr, Smith, & Krakauer, 2010). Among these, everybody learns by experience to expect a temporal match between a movement and its visual feedback as perceived in 1st person. This implies that the brain employs forward models (S. J. Blakemore et al., 1999; S.-J. Blakemore, 2017) for actions generated by the embodied hand, and would in turn engage in error correction when the observed movement does not temporally match the predicted outcome of the executed movement. Sensory prediction errors drive motor adaptation in terms of movement trajectory (Mazzoni & Krakauer, 2006), velocity (Smith, Ghazizadeh, & Shadmehr, 2006; Wagner & Smith, 2008), and timing (Furuya & Soechting, 2010; Maes et al., 2015). Interestingly, the mirror neurons system (Heyes & Catmur, 2022) literature supports the idea that activating inverse and/or forward models via action observation

requires a visuospatial transformation process, to remap the movement into an egocentric frame of reference (Oh et al., 2019). This supports the idea that spatial perspective taking is an embodied cognitive process, in fact the internal emulation of a physical alignment of perspectives (Klaus Kessler & Thomson, 2010). Noteworthy, such transformation comes with a processing cost which is a function of the angular disparity between the observer and the actor (Graf, 1994; Keehner et al., 2006; K. Kessler, 2000; Klaus Kessler & Thomson, 2010; May, 2004; Wraga et al., 2005; Zacks & Michelon, 2005), as if the observer was internally simulating a rotation into the other's point of view. In a dyadic setting, such putative process can be bypassed with the technological means deployed in the present study.

We point out that the sensorimotor account is more parsimonious as compared to the socio-cognitive one, and more conservative in its assumptions. Whereas general theories of brain functioning differ in the assumption that individuals mutually adapt their behaviors based on internal models of the other, they are unified by the shared principle of error minimization (Palmer & Demos, 2021). The validation of either theory is out of the scope of the present work, and we are far from providing conclusive evidence on this matter. Therefore, we lean in favor of the more parsimonious interpretation based on error correction as tenet shared principle. In doing so, we treat dyadic entrainment as property of collective behavior emerging from the interaction, and highlight the effectiveness of the manipulation in tightening the coupling between individuals.

To complement the behavioral findings, we hereby discuss the subjective experiential correlates of mutual embodiment via body-swap. During the visuo-tactile stimulation procedure, participants were capable of systematically discriminating their own hand from the partner's, and showed a general preference for either hand when viewed from a 1st person perspective. These results came with no surprise, since behavioral and neurophysiological evidence supports the existence of a bodily-self recognition mechanism reliant on visual and sensorimotor representations of the hand (Galigani et al., 2021), while the hand orientation with respect to bodily coordinates is a crucial aspect to meet the conditions for embodiment (Pavani et al., 2000). The crucial finding is the interaction effect between factors, showing that the hand of a partner can in fact be integrated in the bodily representation of the self to a greater extent when visually presented in 1st person, as compared to the 2nd person perspective (see Figure *4.3A*). These results confirm the success of our body-swap procedure in eliciting an experiential counterpart to our behavioral findings, suggesting that the putative overlap of self-other representations thought to underpin interpersonal synchronization (Ole Adrian Heggli, Konvalinka, Kringelbach, et al., 2021; Koban et al., 2019) may leverage on the plasticity of bodily representations (Tsakiris, 2010, 2017). In sum, the explicit measures based on self-

reports are coherent with the implicit measures based on attractor dynamic within the dyad. Nevertheless, when asked about sense of ownership and sense of agency experienced during the joint finger-tapping task, participants could very well discriminate between their own's and their partner's hands regardless of the visual perspective (see *Figures 4.3B* and *4.3C*). Sense of ownership subsists as long as sensorimotor congruency is maintained (Kalckert & Ehrsson, 2012; Konstantina Kilteni, Maselli, Kording, & Slater, 2015; Petkova & Ehrsson, 2008), and sense of agency breaks down when the timing of sensory feedback does not match the prediction (Haggard, Clark, & Kalogeras, 2002; Kalckert & Ehrsson, 2012). However, temporal congruency was not constant throughout the task, but rather varying as a function of the drifting metronomes' cycles, which likely led to the breakdown of both illusions of ownership and agency. We should point out that questionnaire items suffer the major flaw of referring to the task as a whole, whereas from our standpoint the most interesting behavioral findings came from an analysis of local dynamics over the course of the interaction.

Finally, we did not find evidence for any association between empathic traits and the proneness to synchronize with the partner when coupled in any visual perspective. This was unexpected, since the link appears to be well documented in the literature (Tzanaki, 2022) and resonates with the idea that attraction to coordinated states is to some extent informative of the most minimal socioemotional connectedness (Marsh et al., 2009). We propose that the negative finding may be attributed to the rigorous competitive nature of the drifting metronomes. Arguably, when task constraints are looser, there is more margin for more empathic participants to intentionally cooperate with each other. On the other hand, the observation that recurrence score systematically dropped as the partners got further away from the cooperation attractor region, shows that our participants consistently attempted to comply with the task and intentionally neglect the partner. We conclude that, when controlling for intention as mediator of dyadic entrainment, the task revealed a dissociation between low-level dyadic entrainment and high-level cognitive empathy. However, among the dimensions of empathy hereby considered (Davis & Others, 1980), empathic concern stood out as predictor for the subjective ratings of embodiment. Specifically, higher emotionality and concern for others predicted a stronger inclination to experience ownership over the hand of another person. This points at the interplay between empathy and the mechanisms underlying embodiment, transferring evidence for such associations from the VR literature (Peck et al., 2018, 2020, 2013; Thériault et al., 2021) to partial body-swap with a real human partner. This particular dimension of trait empathy should not be neglected when adopting the technology for real-world applications, since the evidence suggests it may be a personal variable relevant to the outcome of the manipulation.

Conclusions

The main aim of the present work was to assess whether interpersonal coordination can be facilitated by experimentally inducing the 1st person view of a partner during a rhythmic interaction, as compared to the ecological mode of interaction in 2nd person. Our results support the idea that such manipulation strengthens the coupling between interacting individuals, promoting the cooperation process which facilitates the units of a dyadic system to move together at the collective level of behavior (Rosso, Maes, et al., 2021). From a socio-cognitive viewpoint, we put forward that the dynamic balance between cooperation and competition processes may be underpinned by metastable self-other integration and segregation processes taking place in individual brains during the interaction (Ole Adrian Heggli, Konvalinka, Kringelbach, et al., 2021). The induction of mutual embodiment would then facilitate transitions towards integration and cooperation within the dyad. Whilst this interpretation is plausible, the study does not provide conclusive evidence for a socio-cognitive account. We argue that an explanation based on sensory prediction and adaptation in motor control (Shadmehr et al., 2010) would be more conservative, for it does not make assumptions on the representation of the other in the brain, while accounting for error correction mechanisms leading to dyadic entrainment as emergent property of the interaction.

The major fundamental contribution of our work lies in the observation that dyadic coordination dynamics are subject to the manipulation of visual perspective. Whatever the cognitive mechanism behind, our findings show that we can use it to steer social interactions, supporting joint action by enhancing interpersonal synchronization. Based on our findings, we propose that a technology informed by principles of body-swapping (Petkova & Ehrsson, 2008) has a considerable potential to facilitate interpersonal coordination across a broad range of applications, providing an unprecedented resource in motor training, sports, musical education, and rehabilitation protocols.

Methods

Participants. Forty (N = 40) right-handed participants took part in the study (28 females, 12 males; mean age = 31.42 years, standard deviation = 7.49 years). In order to control for gender bias in the interaction, they were divided in two gender-matched groups and randomly paired in twenty (N = 20) dyads. One dyad was excluded from dyadic analyses due to failure to comply with the instructions. One participant was excluded from the analysis of self-reported ownership due to a technical problem in the video streaming

during the procedure. None of the participants had history of neurological, major medical or psychiatric disorders. All of them declared not to be professional musicians upon recruitment, although some of them had musical experience. None of the participants declared to know the assigned partner from before the experiment. The study was approved by the Ethics Committee of Ghent University (Faculty of Arts and Philosophy) and informed written consent was obtained from each participant, who received a 20€ coupon as compensation for their participation.

Experimental apparatus and procedure. Partners were sitting across the same table, facing each other. During the preparation phase, they were assisted in wearing a black cloth over the whole body and a long red glove over the right hand, with the purpose of rendering the visual scene as neutral as possible and minimizing individual differences related to personal clothing and skin texture. Before proceeding further, each individual participant underwent a period of familiarization with the finger-tapping task. Specifically, an auditory metronome was presented via in-ear plugs and he/she was instructed to tap the right index finger on a circular pad placed on the table. The experimenter showed how tapping was supposed to be performed, so that both partners would adopt a common style during the task. Pink noise was played on the background of the metronomes, with volume adjusted so that each participant could clearly hear the metronome but not the feedback of their own tapping on the pad.

Participants were then equipped with HTC Vive Pro 2 headsets for immersive virtual reality (VR) environments, and underwent the standard calibration procedure as implemented by the manufacturer. Each set was connected to a different computer, running a Unity executable which took video input from a Logitech Brio Ultra HD Pro Business webcam (USB 3.0) and streamed it to the head-mounted display. The setup allowed to present an immersive photorealistic view of the right hand up to the forearm. The hand could be seen in either 1st or 2nd person perspectives, and could either pertain to one's own or to the partner, depending on the experimental condition. Before each condition had taken place, the factor Perspective was manipulated by placing 2 cameras above the shoulder of the participant (1P) or in front of the partner's hand (2P). The factor Ownership was manipulated by simply swapping the USB connection of the cameras to the respective computers, so that participants would perceive their own (Self) or the partner's (Other) hand. The resulting visual scenes can be seen in the details of *Figure 4.1* from all levels of Perspective and Ownership across experimental conditions. Extensive testing prior to the beginning of the study resulted in an average video latency of 96ms for streaming 1080p video at a frame rate of 60Hz. Previous pilots and qualitative interviews with the

participants revealed that the delay was barely perceivable, and in most cases not noticed at all.

In order to collect subjective reports of the sense of ownership over the visually perceived hand, we carried out the following procedure before starting each experimental condition. Both participants were asked to stay relaxed while the experimenter placed their right hand on a cardboard surface on the table, above the assigned tapping pad. They were instructed to watch for one minute the hand that would appear in the head-mounted display shortly thereafter, and to not move their own hand despite whatever would happen in the visual scene. From the moment the experimenter launched the video streaming, participants saw the hand lying still on a cardboard in front of them for 30 seconds (which hand and from which perspective depended on the experimental condition, as illustrated in the details of *Figure 4.1*). For the following 30 seconds, the experimenter applied synchronous touches on the back of the hand of both participants. Next, the visual scene went blank, and the participants were verbally asked the following question: "On a scale from 1 to 5, how much did you feel like the hand that you were seeing belonged to you?". The response was given in silence by raising the fingers of the left hand, in order not to bias the partner and to not induce motor activity in the stimulated hand.

What follows is the description of the 'drifting metronomes' paradigm, as originally described in Rosso et al. (Rosso, Maes, et al., 2021). Each partner was assigned to one pad and instructed to tap on it with the right index finger, synchronizing with an auditory metronome. The two metronomes slightly differed in tempo (1.67Hz and 1.64Hz), whereas the timbre remained the same. With the start of the two metronomes' tracks aligned, the relative phase between metronomes started at 0º and steadily increased in regular steps of 5.6º. A full cycle took 39.008 seconds to be completed (65 and 64 clicks of the faster and slower metronome, respectively). Ten consecutive cycles were performed in each experimental condition. In conditions of informational coupling, participants were instructed to ignore their partner and to tap along with the assigned metronome. Participant's chairs were provided with an armrest, in order to exclude any tactile or proprioceptive coupling due to vibrations of the table resonating with finger taps.

A M-Audio® M-Track 8 soundcard was used to route independent audio channels to each participant via in-ear plugs. The average audio latency from tapping pad to earplug was 17 ms, with a standard deviation of 2 ms. Ableton Live 10 ® was used as main interface for stimuli presentation, with 2 separate MIDI tracks triggering the metronome's audio sample. A Teensy 3.2 microcontroller was used as serial/MIDI hub in the setup: tapping onsets were detected with 1ms resolution using analog input of strain gauge sensors

installed inside the pads, while metronomes onsets were logged using MIDI messages originating from Ableton. Each class of events (metronomes 1 and 2, finger-taps from participants 1 and 2) was retrieved by means of a predefined ID number. Simultaneous EEG recordings were performed from both partners of the dyads during the whole experiment, but such data are not presented in the present paper. Additional data were collected prior and during the experiment. Prior to the experiment, demographical data were collected; the *Edinburgh inventory* (Oldfield, 1971) was administered to assess the right handedness of the participants; the 28-items version of the *Interpersonal Reactivity Index* (IRI) (Davis & Others, 1980) was administered as a self-report of empathy and its subscales. During the breaks between experimental conditions, all participants provided subjective self-reports on different aspects of the task by expressing agreement on a scale from 1 ("Completely disagree") to 7 ("Completely agree") with a custom-made battery of 11 Likert items. Among these, the sense of ownership and sense of agency experienced during the task were measured by asking to rate the respective following items: "I felt like the hand that I was seeing belonged to me", and "I felt like it was me moving the hand that I was seeing". For the sake of conciseness, we reported the analyses of the ones related to the core research questions, namely sense of ownership and sense of agency.

Participants were monitored by the experimenters from behind curtains, where the visual scene of their headsets was visible on two separate screens. Dyad #7 was excluded from the analysis, given that one participant was unable to comply with instructions during Condition #4.

Data Analysis

Pre-processing. Over the course of 10 consecutive metronomes' cycles, 650 and 640 tapping onsets were expected from the partners forming each dyad, for the total duration of 390 seconds. Onsets occurring <350ms from the previous one, were considered false positives and removed, since participants could occasionally push the pad for too long or accidentally lay their hand on it. Out of the whole sample, 59 false positive were removed, corresponding to 0.06% of all data points. The cleaned timeseries were then interpolated with a sine function at 1kHz sampling rate, providing an estimate of the oscillators' positions on its cycle with a temporal resolution of 1ms. The tap preceding the first metronome onset and the last tap following the last metronome onset were included in the interpolation. Afterwards, data points outside the boundaries of the metronomes timeseries were removed. Operationally, the procedure guaranteed equally sized timeseries without loss of data, which was a requirement for the application of joint

recurrence quantification analysis (JRQA; see next paragraph). The modelling of systems of coupled oscillators in the context of joint finger-tapping studies conceptually supports the choice of interpolation (Ole Adrian Heggli et al., 2019; Rosso, Maes, et al., 2021). Finally, timeseries were down-sampled by a factor of 4 to make computation of recurrence plots (RPs) computationally feasible. As shown in our previous report, results of JRQA are robust to the choice of the down-sampling factor.

Phase-space reconstruction. In accordance with Takens' embedding theorem (Takens, 1981), we reconstructed the phase space of individual finger-tapping behaviors. This was done based on time-delayed copies of the input time series u_k, applying an embedding dimension m and a time delay τ.

$$\vec{x}(t) = \vec{x}_i = (u_i, u_{i+\tau}, \dots, u_{i+(m-1)\tau}), \qquad t = i\Delta t$$

where $\vec{x}(t)$ is the vector of reconstructed states in the phase-space at the time t. Optimal parameters for the time-delayed embedding were computed for each participant, for the time course of each single metronome's cycle in all experimental conditions. The resulting mean value was applied to all individual instances. The reason for this approach is that in order to compare the rate of recurrences across conditions at the group level, the embedding procedure must be consistent across participants (e.g., see (Afsar et al., 2018), for an example of parameter selections in a factorial design). We first selected the delay τ as the first local minimum of mutual information index (Fraser & Swinney, 1986) in function of delay. This approach minimized the timeseries self-similarity, extracting nearly orthogonal components and preventing the attractor from folding over itself (Bradley & Kantz, 2015). Mean value of delay resulted to be τ = 7. Next, we determined the number of embedding dimensions with the method of false nearest neighbor (Rulkov et al., 1995). Specifically, we progressively unfolded the time series into higher dimensions until data points did not overlap spuriously, finding an optimal mean embedding of m = 2. Finally, in accordance with the literature, the maximum threshold for counting two neighboring points as recurrent was set at 10% of the maximal phase-space diameter (Marwan et al., 2007).

Joint recurrence plots (JRPs). A recurrence plot $R_{i,j}$ is a square array used to represent and quantify recurrences of states in the phase space of a system (Eckmann et al., 1995). For every point of the phase space trajectory:

$$\vec{x_i}\ (i = 1, \ldots, N; N = n - (m\text{-}1)\ \tau)$$

we tested whether it was close to another point of the trajectory $\vec{x_j}$ based on a neighborhood threshold. Individual recurrence plots were computed as follows:

$$R_{i,j}(\varepsilon) = \Theta(\varepsilon - \|x_i - x_j\|)$$

where ε is the neighborhood threshold, ‖ · ‖ is the Euclidean norm, representing the distance between two vectors, and Θ is the Heaviside step function. A square matrix was returned from each phase-space, containing 1s for all the instances where the distance ‖ · ‖ was smaller than the threshold ε, and 0s for remaining elements. A joint recurrence plot (JRP) was computed for each dyad by pair-wise overlapping partners' individual RPs, and keeping 1s only the instances where both plots contain a recurrence. Each JRP is in fact the Hadamard product of the recurrence plot of the first participant and the recurrence plot of the second participant. Computation of JRPs was carried out using the *crp toolbox* for Matlab ® (Marwan et al., 2007).

The 10 trials (i.e., the metronomes' cycles) of each experimental condition were aggregated by summing the respective JRPs of each trial. This resulted in a 2-D matrix for which every entry contained the amount of recurrences occurring in the corresponding region of the cycle, across all trials. Finally, a 1-D vector recurrence scores was obtained by looping over the columns of the matrix and summing the counts contained in the rows. This vector represents a density measure of the instances of coupled behavior over the course of the metronomes' cycle. The scale of these recurrence scores depend on the size of the JRPs and in turn on the embedding procedure, which makes it necessary to set the same parameters on the whole sample. In order to improve signal-to-noise ratio and avoid over-sampling in view of our statistical model, the resulting timeseries were reduced to 64 bins by averaging the recurrence score for equally sized, consecutive time periods. For this segmentation, interval size was equal to the slower metronome's increments, as they provided a regular subdivision intrinsic to the experimental trials. All processing steps

presented were carried out in Matlab ®. Our approach was preferred over the version for JRQA based on moving windows, for the latter would act as a low-pass filter on our timeseries and hinder the interpretation of our results. Specifically, a moving window result in a phase distortion of the timeseries dependent on window size, and is thus not reliable in detecting attractor points over the attractor landscape. Since the procedure hereby described reproduced exactly the steps in Rosso et al. (Rosso, Maes, et al., 2021), the content of the present paragraph is taken from the original work with the consent of the authors. Values for the embedding dimension and delay do differ from the original work, because they were optimized for the present dataset.

Statistical models. The recurrence score was used as response variable in a mixed-effects model with Coupling and Perspective as factors, and Time as a continuous predictor expressed with the indexes of the metronome's steps (from 1 to 64). Given the non-linear time-course observed in coupled conditions, we adopted the method of orthogonal polynomials (Mirman, 2017) including linear and quadratic functions of Time into our model. Dyads and interactions between Dyads and the factors were modelled as random effects on all polynomial terms, to account for the individual variability in synchronization skills and individual susceptibility to coupling across the experimental manipulations. The random effects structure was used in order to minimize false alarm rates without substantial loss of power (Barr et al., 2013). Informed by our previous study and by the inspection of empirical curves from the present dataset, we limited the polynomial model to the 2nd order as the most parsimonious solution. In this analysis framework, the intercept is considered a 'zero-order' polynomial, as it exhibits zero changes in any direction. Significant changes of direction indicate modulation by the temporal structure of the task. This allowed us to quantify the influence of attractor points, as the dyad deviated from the horizontal trajectory transitioning over expected critical regions. The formula of the full model is the following:

$$Recurrence \sim (Time + Time^2) * Coupling * Perspective + (Time + Time^2 \mid Dyad) + (Time + Time^2 \mid Dyad{:}Coupling{:}Perspective)$$

Aligned rank transform (ART) ANOVA (Wobbrock et al., 2011) was used to test the 2-way interaction between factors on the ratings of sense of ownership over the visually perceived hand, which as an ordinal response variable does not conform to the assumptions of a parametric factorial ANOVA. The same model was fit to the ratings of

sense of ownership and sense of agency referring to the joint finger tapping task. The formulas of the ART ANOVA models are the following:

$$Ownership \sim Coupling * Perspective$$

$$Ownership\_task \sim Coupling * Perspective$$

$$Agency\_task \sim Coupling * Perspective$$

The 3-way interactions of the IRI empathy subscales (Davis & Others, 1980) with the factors were tested by fitting separate mixed-effects linear models for every subscale on the synchronization consistency (with the assigned metronome) and on the ratings of ownership. Subjects were modelled as random effects. The formulas of the two linear models are the following:

$$R \sim Coupling * Perspective * IRI_{subscale} + (1|Subject)$$

$$Ownership \sim Coupling * Perspective * IRI_{subscale} + (1|Subject)$$

Statistical analyses were carried out in R (version 4.0.3). *lme4* (Bates et al., 2014) and ARTool (Kay, Kay, & Wobbrock, 2020) packages were used for model fitting.

Data and code availability

Data and code have been deposited at Mendeley and are publicly available as of the date of publication (https://data.mendeley.com/datasets/24njbrmyjj/1).

DOI: *10.17632/24njbrmyjj.1*

Acknowledgments

The present study was funded by Bijzonder Onderzoeksfonds (BOF) from Ghent University (Belgium), in the context of a joint-PhD project with the University of Lille (France) (I-SITE ULNE program; grant number: 01D21819). The authors are grateful to Ivan Schepers for building the hardware of the finger-tapping device, to Kevin Smink for the illustration of the experimental design, and to Canan Nuran Gener for her precious help in collecting the data.

Author contributions

Conceptualization, M.R.; Methodology, M.R. and B.v.K.; Software, M.R. and B.v.K.; Validation, M.R. and B.v.K.; Formal analysis, M.R.; Investigation, M.R. and B.v.K.; Resources, M.R. and B.v.K.; Data Curation, M.R. and B.v.K.; Writing – Original Draft, M.R.; Writing – Review and Editing, B.v.K., P.M., M.L.; Visualization, M.R.; Supervision, P.M. and M.L.; Project Administration, M.R.; Funding Acquisition, M.R. and M.L.

Declaration of interests

The authors declare no competing interests

5

Mutual beta power modulation in dyadic entrainment

Mattia Rosso [A,B], Ole A. Heggli [C], Pieter J. Maes [A], Peter Vuust [C], Marc Leman [A].

[A.] IPEM - Institute for Systematic Musicology; Ghent University; Ghent, Flanders, 9000; Belgium.

[B.] PSITEC - Psychologie: Interactions, Temps, Emotions, Cognition - ULR 4072; University of Lille; Lille, Hauts-de-France, 59650; France.

[C.] Center for Music in the Brain - Aarhus University; Universitetsbyen 3 – Building 1710; 8000 Aarhus C; Denmark

Rosso, M., Heggli, O. A., Maes, P. J., Vuust, P., & Leman, M. (2022). Mutual beta power modulation in dyadic entrainment. NeuroImage, 257, 119326.

Introduction

The capability of predicting each other's actions is foundational to human interaction (Chris D. Frith, 2007). Without that, we might imagine a social world wherein individuals are doomed to act and react at the wrong time, making interpersonal synergies burdensome at best. In such a world, even common routines such as taking turns in a conversation would be extremely demanding. Playing music, sports or games together would be impossible. All these types of joint action crucially rely on prediction and anticipation, as reaction alone is insufficient to explain the vast spectrum of complexity covered by human interactions (Sebanz et al., 2006; Sebanz & Knoblich, 2009).

Recently, important knowledge has been obtained from studies using minimalistic forms of joint action such as finger tapping (for a review, see Bruno H. Repp & Su, 2013). These types of reduced paradigms offer a highly controlled, yet informative, way of investigating motor control in minimal social interaction (Konvalinka et al., 2010; Novembre, Knoblich, Dunne, & Keller, 2017). Behavioral evidence from joint action experiments have led to the proposal that mutual predictions and shared predictive models may constitute the cognitive machinery behind interpersonal synchronization (Ole A. Heggli et al., 2019; Koban et al., 2019; Konvalinka et al., 2010; Pecenka & Keller, 2011; Rosso, Maes, et al., 2021). Although the behavioral phenomenon could be explained from the perspective of dynamical systems, without assuming that individuals hold internal models of the environment (Demos et al., 2019; Roman et al., 2019; Stepp & Turvey, 2010; Washburn et al., 2017), its underlying neural mechanisms are yet to be mapped (Palmer & Demos, 2021). To this date, this is a topic under active investigation and the main subject of the present work.

A promising framework for investigating the neural mechanisms of interpersonal interactions comes from predictive accounts of brain functioning, which in the past decades have been establishing themselves as general theories of cognition (Clark, 2013; K. Friston, 2005, 2010; K. Friston et al., 2011). The influence of such framework has expanded from the study of the social brain (Chris D. Frith, 2007) to related fields such as interpersonal synchronization (Koban et al., 2019), speech processing (Rimmele, Morillon, Poeppel, & Arnal, 2018) and music cognition (Marc Leman, 2016; Vuust et al., 2018; Vuust & Witek, 2014). The body of evidence suggests that predictive processing underlies common low-level neurophysiological mechanisms across cognitive domains.

One likely neuronal candidate for low-level interpersonal synchronization is beta oscillations of ~20 Hz, which are thought to encode prior representations of the environment (Betti, Della Penna, de Pasquale, & Corbetta, 2021). This narrow frequency

range exhibits functional specificity for top-down processes such as attentional gain (Lee, Whittington, & Kopell, 2013; Nobre & van Ede, 2018; Van Ede, De Lange, Jensen, & Maris, 2011; van Ede, Jensen, & Maris, 2010) and predictions of causal events (van Pelt et al., 2016). Beta power (i.e., the squared magnitude of beta oscillations) is an index of integration and global efficiency of the brain network as a whole, and its fluctuations reflect variations in the level of integration (de Pasquale, Corbetta, Betti, & Della Penna, 2018; De Pasquale, Della Penna, Sporns, Romani, & Corbetta, 2016). These slow dynamics have been put forward as suitable candidate to represent the time-varying structure of external events (Betti et al., 2021). Crucially, it seems to play an important role in the interplay between sensory and motor systems, which ultimately allows us to move in an environment while adaptively responding to its stimuli (Crapse & Sommer, 2008).

Compelling evidence points to beta dynamics as the potential cornerstone for intra- and inter-personal levels of motor control. When humans perform rhythmic behaviors such as finger-tapping, beta activation in sensorimotor regions is periodically modulated as a function of movement cycles (M. Seeber, Scherer, & Müller-Putz, 2016). A similar cyclical pattern of beta modulation has been observed in primary auditory areas in response to rhythmic auditory stimuli, under the condition that these are presented in a predictable temporal structure (T. Fujioka et al., 2009; Takako Fujioka et al., 2015). In the visual domain, beta amplitude in the primary motor cortex reflects the sensitivity to predictable visual stimuli relevant to a motor task (Saleh, Reimer, Penn, Ojakangas, & Hatsopoulos, 2010). Furthermore, beta suppression is consistently observed in the primary sensorimotor cortex in anticipation of predictable tactile stimuli, and associated with faster behavioral responses to such stimulation (Van Ede et al., 2011). The mechanism is thought to reflect temporal orienting of attention, via a dynamic allocation of neural resources which facilitates perception (Nobre & van Ede, 2018). More recently, it was shown that beta power expresses cerebellar “clocking activity” (L. M. Andersen & Dalal, 2021) which may serve to track environmental inputs and prepare the alignment of motor outputs when the context requires it. Beta dynamics based on power modulation appear to be a common denominator when it comes to predictions, since evidence converges towards a shared timing mechanism which the brain might use to pace behavior to rhythmic stimuli, via action-perception loops based on predictive control.

In social environments, other humans represent exceptionally salient social stimuli, and a unique class of rhythmic affordances for the motor system (Phillips-Silver et al., 2010). Humans have a strong tendency to synchronize their movements with each other, often spontaneously (Oullier et al., 2008; Richardson, Marsh, Isenhower, et al., 2007; Shockley, Santana, & Fowler, 2003) or even against explicit instruction to ignore each other (Rosso, Maes, et al., 2021). Interestingly, the brain decodes the kinematics of visually perceived

movements via modulation of beta power (G. Zhou, Bourguignon, Parkkonen, & Hari, 2016), and evidence from a dual-brain stimulation paradigm suggests that 20 Hz bursts over the motor cortex causally contribute in establishing and maintaining synchronous dyadic behavior (Novembre et al., 2017).

In this study, we hypothesized that the human brain may rely on a shared mechanism based on beta modulation, both to track self-generated behavior and to predict the behavior performed by another individual. In order to test this hypothesis, we analyzed dual-EEG data recorded during an interpersonal synchronization task. In the 'drifting metronomes' paradigm (Rosso, Maes, et al., 2021), dyads of participants were cued with metronomes set at different frequencies, and depending on the condition they perceived each other visually or auditorily. In contrast to most dual-EEG experiments which focus on measures of interbrain synchronization, we here computed a beta power timeseries for each participant as an intrabrain measure (Hamilton, 2021), and coupled it to movement cycles produced by oneself (Self) and by the partner (Other). Whilst there is evidence that some interbrain coupling measures are genuinely related to aspects of social interaction, their interpretation is not always transparent, and their susceptibility to spurious correlations calls for extra caution (Burgess, 2013; Czeszumski et al., 2020). Given a minimalistic joint finger-tapping task and the control over the perceptual coupling between partners, we explicitly modelled the neural dynamic of interest as a function of overt behavior occurring in the interaction. In presence of mutual coupling, we expected to observe cyclical beta power modulation in each individual, corresponding to movements performed by Self and Other (*Figure 5.1*).

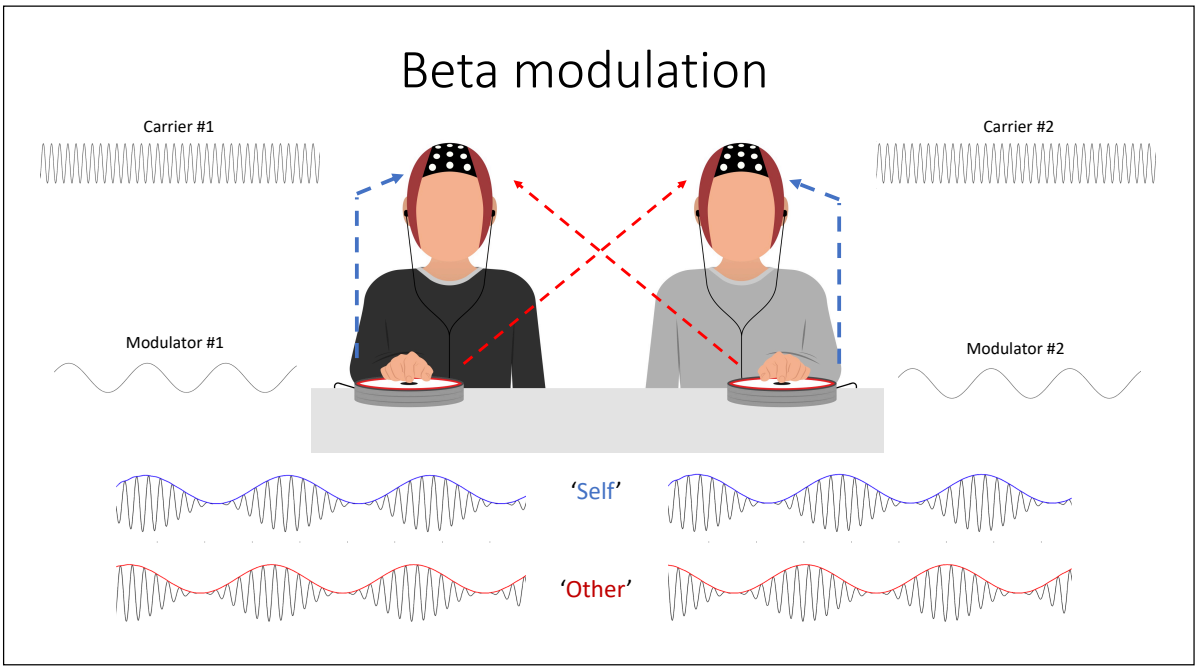


Figure 5. 1. Beta modulation by Self and by Other. The figure shows a schematic representation of two participants engaged in a joint finger-tapping task, while undergoing dual-EEG recording. For both of them, a high-frequency component of interest was extracted from the neural timeseries (the 'carrier', at ~20 Hz), and a low-frequency signal from the behavioral timeseries (the 'modulator', at the tapping frequency). The dashed arrows represent two instances of beta modulation operationalized in the same setting: the blue one represents modulation by self-generated tapping cycles, whereas the red one represents modulation by other-generated tapping cycles. The plots on the bottom show the envelope of the neural signal as amplitude-modulated by both Self (blue) and Other (red). Whilst we expected to find a periodic modulation by Self due to the finger-tapping task per se, we expected to observe modulation by Other exclusively in conditions where the partners were able to perceive each other's movement. The signals hereby represented are simulated and shown for illustrative purposes only.

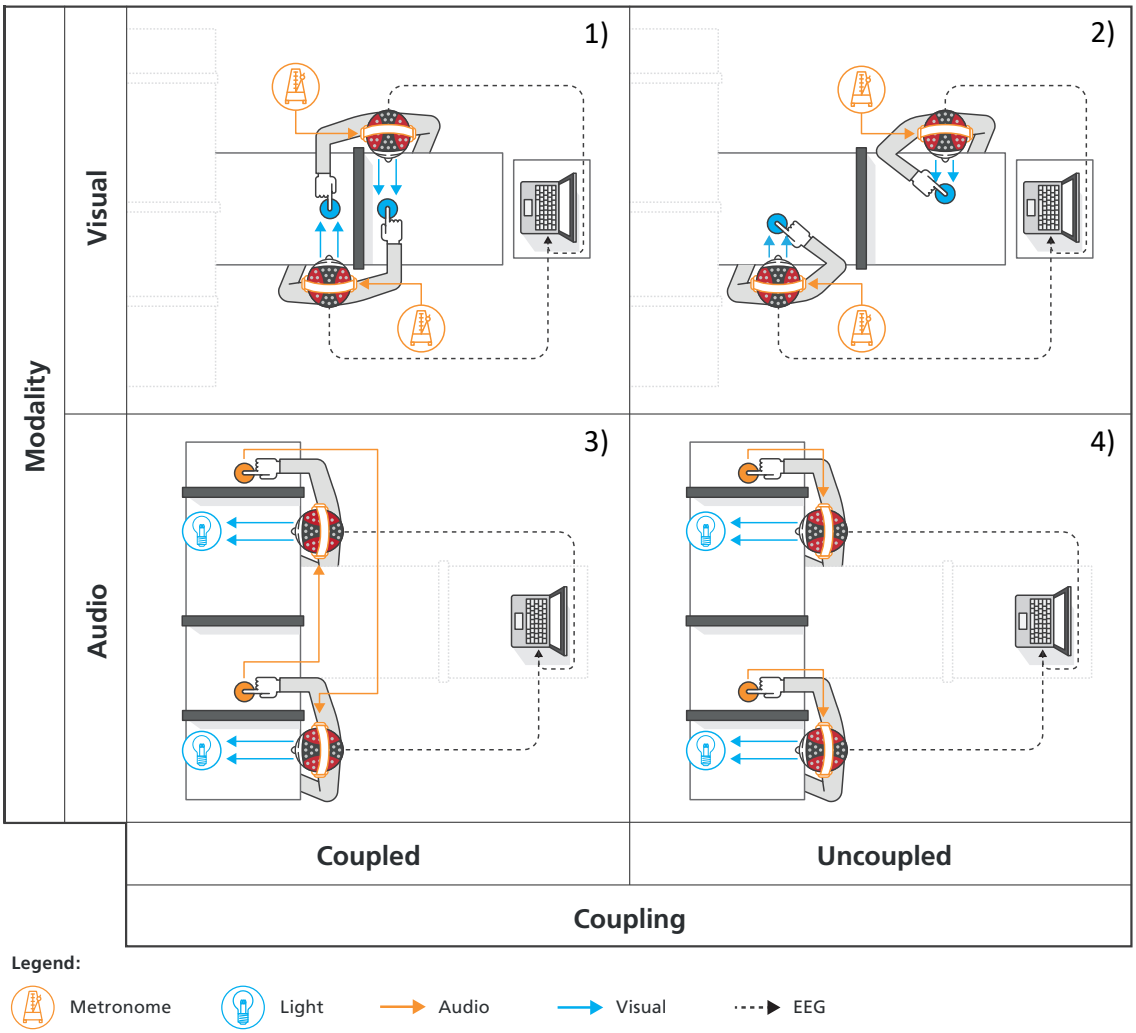


Figure 5. 2. Experimental design. The experiment consisted of a 2 x 2 factorial structure: Modality (Visual, Auditory) x Coupling (Coupled, Uncoupled). The design resulted in the following conditions. **1.** ***Visually Coupled.*** Participants were instructed to synchronize their finger-tapping with an auditory metronome, while looking at the partner's hand. The view of their own hand was hidden by a screen placed on the table. **2.** ***Visually Uncoupled (control).*** Participants were instructed to synchronize their finger-tapping with an auditory metronome, while looking at their own hand. The view of the partner's hand is hidden by a screen placed on the table. **3.** ***Auditorily Coupled.*** Participants were instructed to synchronize their finger-tapping with a flickering LED, while hearing the sonification of the partner's tapping. The view of both participants' hands was hidden by screens placed on the table. They were informed that the sounds they would hear were produced by the partner. **4.** ***Auditorily Uncoupled (control).*** Participants were instructed to synchronize their finger-tapping with a flickering LED, while hearing the sonification of their own tapping. The view of both participants' hands was hidden by screens placed on the table. They were informed that the sounds they would hear were produced by themselves. In coupled conditions (1 and 3), participants were clearly instructed to neglect the movements produced by the partner and to synchronize with the assigned metronome at all times. The figure is adapted from (Rosso, Maes, et al., 2021), and includes the representation of the dual-EEG setup.

Materials and Methods

1. Participants

Twenty-eight (N = 28) participants took part in the study (18 females, 10 males; mean age = 29.07 years, std = 5.73 years). All of them were right-handed, without history of neurological or psychiatric disorders. None of them was a professional musician, although some had musical experience. Handedness was assessed by means of the Edinburgh Handedness Inventory (Oldfield, 1971). Upon recruitment, participants were divided by gender and randomly paired. The experiment was approved by the Ethics Committee of Ghent University (Faculty of Arts and Philosophy) and informed written consent was obtained from each participant. Participants received a 15€ voucher as economic compensation for their time.

2. Behavioral Task

The two partners forming a dyad were sitting at the same table facing each other. In order to avoid any tactile or proprioceptive coupling, they were asked to lay their elbow on an armrest at all times and avoid contact with the table. Each partner was assigned to one circular tapping pad and instructed to tap on it with the right index finger, with the goal of synchronizing with a metronome. Depending on the experimental condition, the metronome could either be auditory or visual, and the partner was always perceived in the complementary modality. Each partner was cued with a slightly different metronome tempo (100 BPM and 98.5 BPM, or 1.67 Hz and 1.64 Hz respectively in terms of frequencies), resulting in a pattern of continuous de-phasing. Hence, the metronomes' relative phase gradually increased from 0 to π radians, and then decreased from π to 0 radians over the course of 10 consecutive cycles, for a total duration of 390 seconds. Participants were always instructed to ignore the partner and tap along with the assigned metronome. For more details on the implementation and the rationale of the behavioral task, we refer to (Rosso, Maes, et al., 2021). Setup and experimental design are illustrated in *Figure 5.2*.

Demographic data were collected before the experiment, along with the administration of the Edinburgh Inventory for handedness assessment and the Interpersonal Reactivity Index (IRI) (Davis & Others, 1980). After every experimental condition, participants provided self-reports on different aspects of the task by expressing agreement on a scale from 1 ("Completely disagree") to 7 ("Completely agree") with a custom-made battery of 11 Likert items. The complete battery is provided as *Supplementary material*. From the

whole battery, we analyzed two items which are relevant to the present study: “I found it difficult to keep the tempo of my metronome”, and “I felt in control over my own actions”.

3. Experimental apparatus

Each participant was provided with a circular pad containing a strain gauge pressure sensor, used to detect tapping onsets with a 1 ms resolution. This was connected to a Teensy 3.2 microcontroller, which worked as serial/MIDI hub to log the data and communicate with the other devices. A detailed description of the setup, behavioral data acquisition, stimuli generation and audio routing can be found in (Rosso, Maes, et al., 2021). At the beginning of each metronomes’ cycle, a TTL trigger was sent from the Teensy microcontroller to the EEG amplifier via BNC connection, guaranteeing the synchrony of behavioral and neural timeseries. The EEG signals were recorded with an ANT-Neuro eego™mylab system at a sampling rate of 1 kHz. Each participant was equipped with an EEG headset (64-channel waveguard™original with Ag/AgCl electrodes).

4. Behavioral data processing

Since participants could occasionally produce an artificial ‘double tap’ by pushing the pad for too long, we removed events whenever an onset followed the previous one by less than 350 ms. Subsequently, we calculated the phase time-series from every participant’s taps by interpolating the onsets as a ramp wave at 1 kHz sampling rate and scaled to 2π radians. Via this procedure, we explicitly modelled the participants as oscillators (Ole Adrian Heggli et al., 2019) and provided an estimate of their phase with a temporal resolution of 1 ms. Data processing was carried out in Matlab®.

5. Dual-EEG data acquisition

Each participant was equipped with a 64-channels *waveguard™original* EEG headset (10-10 system, with Ag/AgCl electrodes). Two ANT-Neuro *eego™mylab* systems were connected in cascade to synchronize the recordings, which were performed at a sampling rate of 1 kHz. Each pair of headsets shared a common ground, and “CPz_partner1” was used as common reference electrode for both of the partners. Impedances were monitored in the eego™ software environment and kept below 20 kΩ.

6. EEG pre-processing

The pre-processing pipeline was written using functions from the *Fieldtrip* toolbox for *Matlab* (*MathWorks Inc, USA*). Bad channels were identified by visually inspecting the raw timeseries and the distribution of variance across channels. For every dyad, the 128-channels recordings were divided in 2, and re-referenced to the electrodes "CPz_partner1" and "CPz_partner2", respectively. Only after the rejection of bad channels, we re-referenced the two recordings to the average of all the respective 64 electrodes, to avoid noise leakage into the common average. An average of 1.14 bad channels per participant were removed (std = 2.07). A high-pass sixth-order Butterworth filter with a 1 Hz cut-off was applied to the raw recordings to remove slow drifts, a conservative threshold given the length of the recordings. A low-pass sixth-order Butterworth filter with 45 Hz cut-off was applied to remove high-frequency muscular activity. A fourth-order notch filter centered at 50 Hz was applied to remove power-line noise up to the 3rd harmonic.

Blinks and eye-movement artifacts were removed by means of visual inspection of topographical maps and component activation timeseries. For this purpose, we ran independent component analysis (ICA) using the 'runica' algorithm as implemented in Fieldtrip, excluding the reference 'CPz' and the bad channels' timeseries from the input matrix. Only those components which exhibited the stereotypical frontal distribution generated by blinks and lateral eye movements were removed. The selection was limited to few unambiguous components for the sake of replicability. A minimum of 1 and a maximum of 3 components were removed for every participant. The dataset was inspected prior to ICA decomposition and the following ICA back-projection. Special attention was given to frontal clusters of electrodes, where the activation of the artifactual components was maximal. Rejected bad channels were reconstructed after artifact removal, by computing a weighted average of activity from neighboring electrodes. No segmentation in epochs was performed on the recordings, so that every experimental condition was treated as a continuous experimental run.

7. EEG source separation

In order to reduce the dimensionality of the data, we designed a spatial filter using generalized eigendecomposition (GED). The technique allowed us to perform a hypothesis-driven source separation on the multivariate EEG signal, guided by a certain criterion (Michael X. Cohen & Gulbinaite, 2017; Rosso, Leman, et al., 2021). In our case, we wanted to find the weighted combination of electrodes which best separated beta

activity (~20 Hz) from the broadband signal, so that we could disentangle spatial patterns of dominant beta band activation. The procedure hereby described was applied to all participants, separately for all experimental conditions.

Whereas in its original formulation the technique makes no previous assumption on scalp sources (Michael X. Cohen & Gulbinaite, 2017), our target component was computed as the weighted average of the 37 channels located behind the frontocentral 'FC' line - mastoids excluded (see also (Rosso, Leman, et al., 2021)). The rationale for this macro-selection was to have full coverage of centroparietal, temporal and occipital regions where beta modulation was expected to be maximal (G. Zhou et al., 2016). For consistency, analyses were replicated including electrodes from the whole scalp (mastoids excluded) and yielded the same final results on beta modulation. However, given the poor consistency of the spatial activation patterns observed across participants, we decided to optimize the source separation by setting a spatial constrain, in order to present here physiologically interpretable results. The improvement might be explained by residual noise in frontal channels caused by eye-related artifacts, or by more prominent beta activity in frontal generators for some of the participants.

The vector of weights W was calculated by solving the following eigenequation:

$$R^{-1}SW = \Lambda W$$

where S is the covariance matrix calculated from the narrow-band filtered signal; R is the reference covariance matrix calculated from the broad-band signal; Λ is a set of eigenvalues. GED identifies eigenvectors W that best separate the signal ('S') covariance from the reference ('R') covariance matrix. The eigenvector with the largest eigenvalue was used as spatial filter. Raw channel data were then multiplied by the eigenvector, to produce the single timeseries of our target beta component.

The S ('signal') covariance matrix was computed from the narrow-band multivariate signal. Using the *firls* Matlab function, we designed a finite impulse response (FIR) filter, with a 18-22 Hz range and 15% slope to avoid edge artifacts in the time domain. In order to minimize the bias towards the center frequency (20 Hz), we opted for a plateau-shaped filter to increasingly suppress frequencies as a linear function of the distance from the range boundaries, while leaving the central range unaffected. The filter resulted in maximum suppression below ~15 Hz and above ~25 Hz, providing good coverage for a big portion of the lower beta range (for a visual representation of the frequency response, see

Supplementary material 6.A). The multivariate signal was filtered in forward and reverse directions to achieve zero phase distortion. The R ('reference') covariance matrix was computed from the broadband multivariate signal, containing frequencies from 1 to 45 Hz. In order to compute the respective covariance matrices from the broad- and narrow-band signals, we used the onset timing of the finger-taps performed by the participant to define time-windows from -100 ms to 500 ms around the events. For every recording, 645 covariance matrices were expected on average, and we removed the outliers whose Euclidean distance from the grand-average covariance matrix exceeded the 2.23 z-score (i.e., corresponding to a probability of 0.013). With this approach, grand-average S and R covariance matrices did not include occasional burst of artifactual activity occurring over the long recording.

It should be noted that, in our experiment, stimulations of different nature were presented to the participants across different conditions. This fact represented for us a strong argument for computing the covariance matrices separately for each condition. Different percepts imply different perceptual processes and, in turn, different activation patterns should be expected in the participants' brains. A clean separation of the beta component within every experimental condition is a necessary condition to observe the modulation of interest, and ultimately compare the effect across conditions. For consistency, we replicated the analyses by computing S and R covariance matrices including data from all conditions, before applying the resulting eigenvectors to the broadband data from each individual condition. The latter approach had a detrimental impact on our analyses, disrupting even the most trivial effects. As a note for the reader, we conclude that the selection and aggregation of data should always be informed by the particular experimental design when implementing a spatial filter via GED.

8. Power-by-phase modulation

The extracted component was filtered in the 18-22 Hz beta range and Hilbert-transformed. Beta power was computed as the squared magnitude of the analytic signal, and extreme values deviating from the mean by more than 3 standard deviations were considered outliers and removed. For each participant, beta power modulation was computed as a function of the self-produced ('Self') and the partner's ('Other') finger-tapping cycles. Finger-tapping phase timeseries were divided into 36 bins (bin size = 10°) (G. Zhou et al., 2016) and the median power values falling within the same bins were computed. From this procedure, we obtained the beta power curves as a function of the movement cycles, computed over the total amount of finger-taps (645 events were

expected on average from each participant). The whole processing pipeline is illustrated in *Figure 5.3*.

9. Statistical model

Our modelling was motivated by the observation that beta power modulation could be well approximated by a sinusoidal function of movement phase, bounded within the unit cycle across participants. The best fitting sinewave was estimated using the *sineFit* Matlab® function (Seibold, 2021), for each participant in each experimental condition, and the sine amplitude provided us with a measure of the beta modulation strength. Sine amplitudes were log-transformed for better model fitting, and used as response variable in the following mixed-effects model:

$$Beta\_modulation \sim Who*Coupling*Modality + (1 \mid Subject:Dyad)$$

where Who (Self, Other), Coupling (Coupled, Uncoupled) and Modality (Visual, Auditory) were modelled as 2-levels categorical fixed effects. Factors were releveled such that 'Other' and 'Uncoupled' would provide the baseline intercept for the model: since the modulation by the other's movements is expected to be null in absence of coupling between the partners, the curve was expected to result in a flat line. The interaction between individual Subjects and the respective dyad was modelled as random effect.
As supplementary analysis, the same model was fit to alpha modulation, as a control frequency band to support the specificity of the effects for beta. Such frequency band was chosen in light of the importance traditionally attributed to in the context of perceptually-mediated interpersonal coordination (Tognoli & Kelso, 2015; Tognoli, Lagarde, DeGuzman, & Kelso, 2007). The response variable was computed with the same procedure described up to this point, with the only difference that the broadband signal was filtered in the 8-12 Hz range (slope = 15%) to extract the component of interest.

$$Alpha\_modulation \sim Who*Coupling*Modality + (1 \mid Subject:Dyad)$$

In order to gain further insight into our results, an ordinal logistic regression model was fitted to two 1-7 Likert items of interest, namely "I found it difficult to keep the tempo of my metronome" and "I felt in control over my own actions":

$$Perceived\_Difficulty \sim Coupling * Modality$$

*Sense_of_Control ~ Coupling * Modality*

Behavioral and EEG data analyses were entirely carried out in Matlab® (version R2019a). Statistical analyses were carried out in R (version 4.0.3). Data and scripts are available upon request to the authors with a formal data sharing agreement, in line with the conditions of the local ethics committee which approved the present study.

Results

The mixed effect model showed a significant main effect of Who (*Estimate* = 0.679; *SE* = 0.180; *t* = 3.768; *p* < 0.001), indicating that the modulation of beta power by self-generated movements was significantly stronger compared to the modulation by the partner's movements.

We found a significant main effect of Coupling (*Estimate* = 0.398; *SE* = 0.180; *t* = 2.208; *p* < 0.05). This indicates that, when partners were coupled visually or auditorily, beta power was modulated to a greater extent by both the self-generated and the perceived movements, as compared to the uncoupled baseline conditions. We furthermore observed a close-to-significant interaction between Who and Coupling (*Estimate* = -0.462; *SE* = 0.255; *t* = -1.813; *p* = 0.07) suggesting that, when coupled, the modulation by the perceived movements increased more than the modulation by the self-generated movements. We found no significant effects of Modality, nor interactions with the other factors. The model passed the Shapiro-Wilk test for normal distribution of the residuals (*p* = 0.25). The results are visually represented in *Figure 5.4*.

In the questionnaires, we found significant main effects of Coupling (*Estimate* = 0.750; *SE* = 0.283; *t* = 2.650; *p* < 0.01) and Modality (*Estimate* = -1.342; *SE* = 0.297; *t* = -4.514; *p* < 0.001) on the perceived Difficulty of the task, along with significant main effects of Coupling (*Estimate* = -0.848; *SE* = 0.283; *t* = -2.994; *p* < 0.01) and Modality (*Estimate* = 0.833; *SE* = 0.298; *t* = 2.791; *p* < 0.01) on the perceived Sense of Control over the performed finger-tapping. Hence, whenever the partner was visually or auditorily perceived, participants perceived the task as more difficult and reported a reduced sense of control over their own movements. This was equally true in all conditions where participants where paced visually. All model summaries are reported in *Table 5.1*.
No significant effects were found for alpha modulation, with the exception of the main effect of Who (*Estimate* = 0.360; *SE* = 0.128; *t* = 2.814; *p* < 0.01). This indicates that, whilst alpha power is significantly modulated by Self-generated movements, there is no evidence

that Coupling between participants results in mutual alpha modulation. This further supports the evidence that the phenomenon hereby reported is specific to beta. Figures and table related to this supplementary analysis can be found in the *Supplementary materials*.

Finally, by multiplying the selected eigenvector by the covariance matrix (w^TS), we obtained an estimation of the underlying source projecting onto the electrodes (Michael X. Cohen, 2022). The resulting spatial patterns of our beta component were computed for every participant in every experimental condition. *Figure 5.5* shows the grand-average pattern and the contrast between coupled and uncoupled conditions, separately for every sensory modality. In contrast with eigenvectors, spatial patterns are physiologically meaningful and can be visually interpreted with due caution.

We mention here as a reminder that both the spatial patterns and the timeseries analyzed in the present work refer to the 1st component extracted via GED (i.e., via the eigenvector associated to the highest eigenvalue), which was meant to extract the most prominent beta activity. Whilst such choice represents a common practice, there is no certainty that the 1st component would be the most beta-modulated for all participants: beta power and beta power modulation are in principle dissociable, since high power does not imply strong modulation, and vice-versa. For the sake of completeness, we verified that the patterns of beta modulation across conditions are observable for 2nd and 3rd components at the group level, and included in the *Supplementary materials* the spatial patterns of the top 3 components for every participant in every experimental condition.

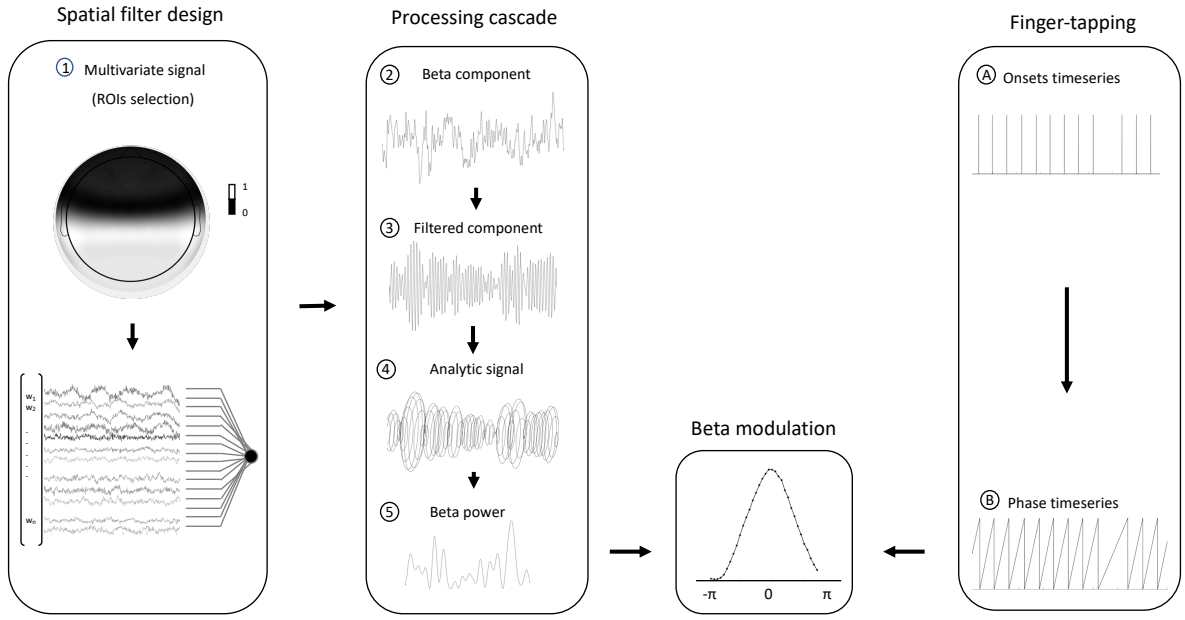


Figure 5. 3. Analyses pipeline. The present pipeline illustrates the steps which led to computing the beta modulation for each individual participant. The steps identified by numbers (on the left) refer to EEG data and start from the completion of pre-processing, whereas the steps identified by letters (on the right) refer to behavioral data and start from the completion of false positives removal. Although the following explanation starts from the processing of EEG data, the two modules may be run in parallel. Generalized eigendecomposition (GED) was performed on a broad set of regions of interest (ROIs) to best separate narrow-band ~20 Hz beta activity from broad-band activity. The vector of weights w associated with the highest eigenvalue was used as spatial filter: the dimensionality of the data was reduced to a single timeseries by computing the weighted average of the 37 channels located behind the frontocentral line (1). The weights of the excluded channels were set to 0. The resulting beta component (2) was then narrow-band filtered with a plateau-shaped FIR filter centered at 20 Hz, in order to exclude the contribution of broad-band activity from the envelope (frequency range = 18-22 Hz; slope = 15%). Finally, we Hilbert-transformed the filtered component (3) to produce the analytic signal (4), and computed the beta power timeseries as its magnitude squared (5). In parallel, we linearly interpolated the discrete finger-tapping onsets (A) to produce an estimate of their phase-timeseries (B). Finger-tapping phase timeseries were divided into 36 bins (bin size = 10°) (G. Zhou et al., 2016), and the EEG power values falling within the same phase bins were averaged together. Beta power is represented as a function of the average finger-tapping cycle, computed for all participants from the totality of their tapping onsets.

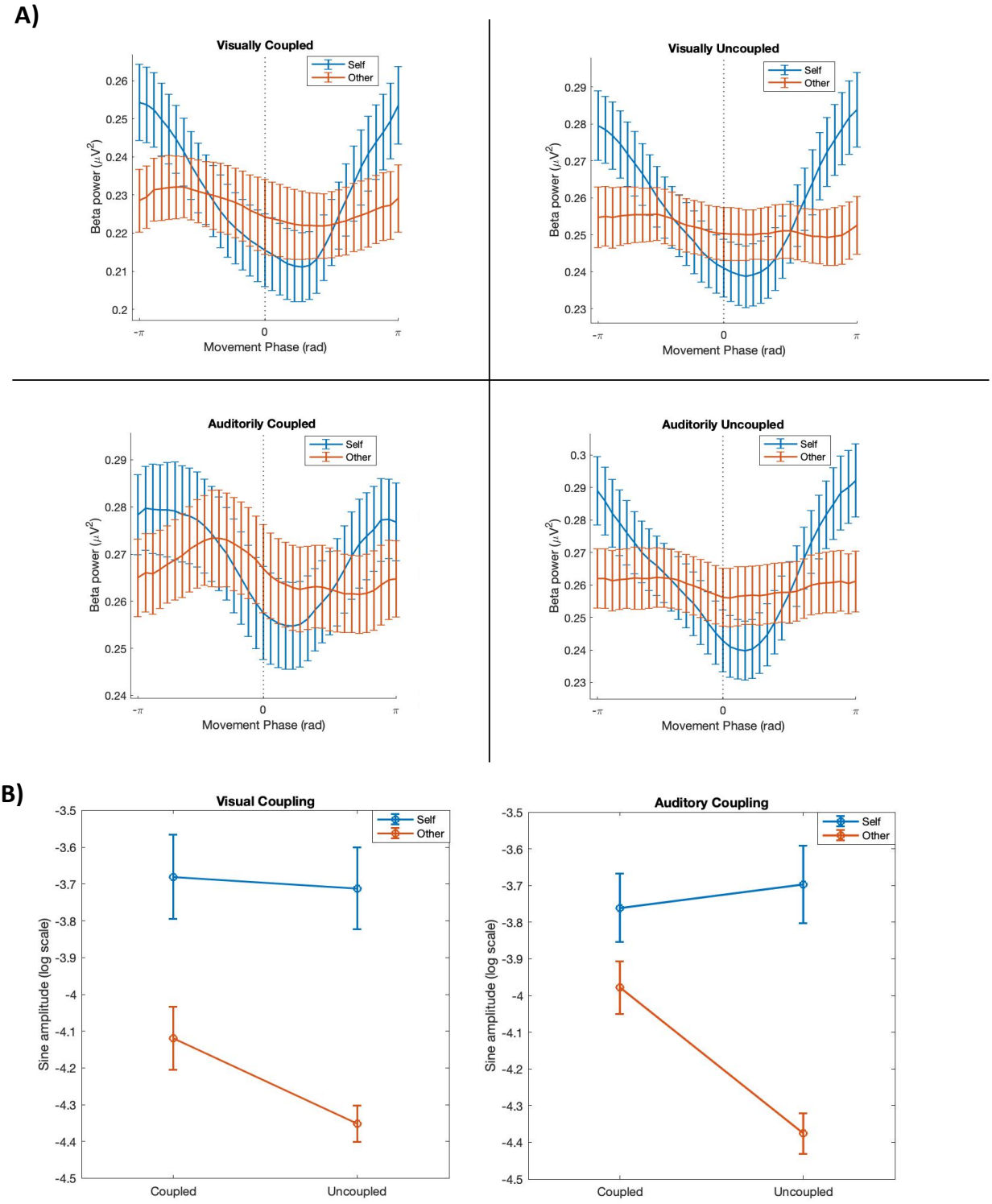


Figure 5. 4. Beta modulation. *A) The grand-average beta-power curve (N = 28) as a function of the tapping cycles* is shown for each experimental condition. In order to highlight the modulation of interest, the y-axis range is normalized for the maximum and minimum values of the curves in each condition, and horizontal dashed lines intercept the average value of the curves. Blue curves represent the grand-average power as a function of Self-

generated finger-tapping cycles, whereas the red curves represent the modulation by the Other. Whilst the former exhibits quasi-sinusoidal course around the intercept in all conditions, the latter does so in coupled conditions only (although to a lesser extent). The perception of the partner's movements is necessary condition for beta power to be modulated. For visualization purposes, the curves were smoothed with a moving average (span = 7 phase bins). Error bars represent the standard error of the mean (SEM). *B) The amplitude of the best-fitting sinewave* was used as a measure of modulation, and the grand-average is hereby shown. As confirmed by our statistical model, modulation by Self is consistently higher across experimental conditions (main effect of Who), and coupled conditions lead to a global increase of modulation (main effect of Coupling). The modulation by Self in the auditorily coupled condition represents an exception, although the decrease from the baseline was not strong enough to reach significance. In fact, we found no significant effect of Modality, nor interactions with the other factors. Ultimately, the figure highlights how coupling leads to a greater increase in the modulation by Other from the uncoupled baseline, as compared to the modulation by Self (interaction effect between Who and Coupling). Error bars represent the SEM.

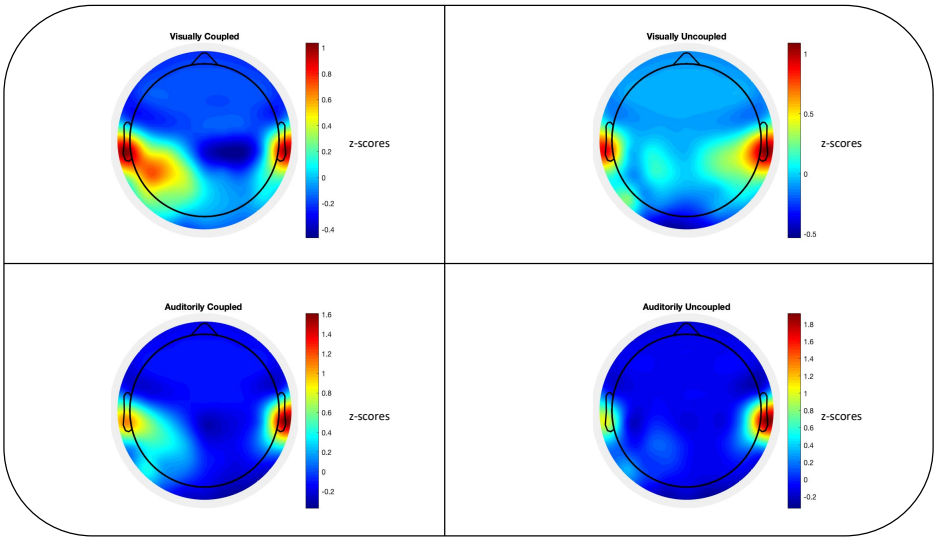
	Modulation amplitude (N = 28)			
Predictors	*Estimate*	*SE*	*t value*	*p*
Who	0.679 ***	0.180	3.768	**< 0.001**
Coupling	0.398 *	0.180	2.208	**0.027**
Modality	0.024	0.180	0.134	0.893
Who:Coupling	-0.462	0.255	-1.813	**0.070**
Who:Modality	-0.039	0.255	-0.153	0.879
Coupling:Modality	-0.165	0.255	-0.649	0.516
Who:Coupling:Modality	0.260	0.360	0.723	0.470
	** p < 0.05*	*** p < 0.01*	**** p < 0.001*	

	Perceived difficulty (N = 28)			
Predictors	*Estimate*	*SE*	*t value*	*p*
Coupling	0.750 **	0.283	2.650	**0.008**
Modality	-1.342 ***	0.297	-4.515	**< 0.001**
Coupling:Modality	-0.225	0.402	-0.560	0.576
	** p < 0.05*	*** p < 0.01*	**** p < 0.001*	

	Sense of control (N = 28)			
Predictors	*Estimate*	*SE*	*t value*	*p*
Coupling	-0.848 **	0.283	-2.994	**0.003**
Modality	0.833 **	0.298	2.791	**0.005**
Coupling:Modality	0.311	0.407	0.763	0.445
	** p < 0.05*	*** p < 0.01*	**** p < 0.001*	

Table 5. 1. Models' summaries All significant effects are marked with an asterisk, and the associated p-values are highlighted in bold. For the beta modulation amplitude, we reported a significant main effect of Who, a significant main effect of Coupling and a close-to-significant interaction effect between Who and Coupling. No significant effects of Modality were found. The directionality of the effects was assessed by setting 'Other' (Who) and 'Uncoupled' (Coupling) as baselines for the contrasts across factors' levels. The reason is that in uncoupled conditions, the modulation by the partner was expected to result in a flat line, providing the baseline level to test the significance of the modulation amplitude. For the ratings of perceived difficulty and sense of control, our ordinal logistic regression model revealed significant main effects of Coupling and Modality. The effect of Modality on the subjective ratings was not discussed in the main text, since it refers specifically to the nature of the stimuli and does not appear to have a relevant link with beta modulation.

A)



B)

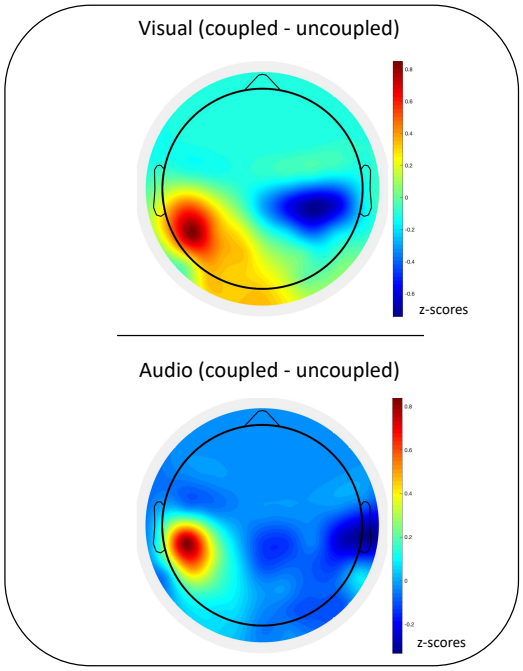


Figure 5. 5. Beta component's spatial patterns. A) The beta activation coefficients were z-score normalized within every individual participant, and the grand-average (N = 28) was computed for every experimental condition. Bilateral temporal activation stands out as the common patterns across conditions (maximum at 'T7' and 'T8'), whereas a left centro-parietal cluster emerged selectively in coupled conditions regardless of the sensory modality. B) The coupling-specific cluster is highlighted by subtracting the Uncoupled conditions maps from the Coupled condition maps (maximum at 'CP5' for visual coupling, at 'C5' for auditory coupling). As stated in the main text, we refrain from drawing conclusions on the neural networks underlying the activation observed at the sensor-level. Our discussion is limited to the evidence of the coupling-specific pattern, which is consistent across visual and auditory modalities.

Discussion

With the present work, we provided evidence for a common neural mechanism used by the brain to control rhythmic behavior and to track the rhythmic behavior produced by a partner during interpersonal synchronization. Specifically, we uncovered brain dynamics of interest in a narrow beta range centered at 20 Hz, consistent with converging evidence of its pivotal role in motor control, timing and top-down predictive processing. Using a joint finger-tapping paradigm (Rosso, Maes, et al., 2021), our results showed sinusoidal modulation of beta power as a function of both self-generated and perceived finger-tapping cycles.

The observed modulation is consistent with the mutual prediction theory, which proposes that during a social interaction each individual utilizes brain systems to control their own behavior and systems to predict the behavior produced by another individual (Hamilton, 2021). Our study shows that beta modulation may be a marker of this dual function in the context of joint tapping, indicating a common neural mechanism for controlling your own as well as predicting the partner's behavior (Sebanz et al., 2006; Sebanz & Knoblich, 2009). While in animals subsets of neuronal populations are deployed for carrying out these two functions (Kingsbury et al., 2019), the human 'social brain' is thought to use dedicated structures for generating predictions during social interactions (Chris D. Frith, 2007).

At the functional level, oscillations around 20 Hz exhibit some unique time-domain features which make them prone to couple to either internal or external events. Transient bursts of activity in the range occur as a spontaneous probabilistic phenomenon at rest (Feingold, Gibson, DePasquale, & Graybiel, 2015), and operate as a relay function of distal inputs in presence of stimulation (Sherman et al., 2016). The coexistence of sustained fluctuations and transient bursts within the same frequency range has led to a reconsideration of the functional roles of oscillatory activity, and to the proposal that bursting activity might enable transient communication between neuronal populations oscillating around the same center frequency (van Ede, Quinn, Woolrich, & Nobre, 2018). The ubiquity of beta dynamics in motor control is long-known (Bourguignon, Jousmäki, Dalal, Jerbi, & De Tiège, 2019; Neuper & Pfurtscheller, 2001; Pfurtscheller, 1981; Salenius, Schnitzler, Salmelin, Jousmäki, & Hari, 1997; Salmelin & Hari, 1994), which led us to hypothesize that its periodic modulation could play a critical role in action-perception loops underpinning overt synchronization behavior. Motoric sampling routines temporally structure the gathering of perceptual information (Schroeder et al., 2010) throughout the animal kingdom and across sensory modalities (Hatsopoulos & Suminski, 2011; Kleinfeld, Ahissar, & Diamond, 2006; Wachowiak, 2011) by transiently modulating ongoing neural oscillations in sensory areas (Crapse & Sommer, 2008). The production of rhythmic

behaviors results in dynamic fluctuations in network excitability, which operates as a rhythmic mode of attention towards incoming streams of perceptual events (Schroeder & Lakatos, 2009). Remarkably, it has been observed that beta oscillations originating in the primary sensorimotor cortex and directed towards auditory areas encode the temporal selection of relevant auditory features, while overt rhythmic behavior enhances the predictive efficacy of perceptual events (Morillon & Baillet, 2017).

A dynamic attending perspective (Large & Jones, 1999) provides a broader framework for understanding our finding in terms of neural oscillations and their dynamics, offering a complementary view to predictive coding for explaining interpersonal synchronization. According to the theory, expectancies of future events are enabled by internal oscillations (attending rhythms) driven by distal events (external rhythms). The presence of a temporal structure in perceptual stimuli is crucial for the attender to generate accurate predictions of the upcoming events. Although originally formulated as entrainment in the strict sense (Large & Jones, 1999), dynamic attending could be based on a mechanism of amplitude modulation where the energy of high-frequency attending rhythms varies as a function of the phase of a low-frequency external rhythm. *Figure 5.1* shows a schematic of how this putative mechanism could enable mutual predictions in an interacting dyad. The novelty of our approach consists of explicitly coupling beta power to Self-generated and Other-generated movements in the context of the same dyadic task. Within our experimental design, we manipulated the presence and the sensory modality of informational coupling between partners: in absence of coupling, the modulation curve by Other was expected to result in a flat line, providing the baseline for our statistical model. The 'drifting metronomes' procedure guaranteed that the expected uniform distribution of relative phase between the partners would be uniform (Rosso, Maes, et al., 2021), controlling for spurious modulations on the curve shape.

Our results showed that beta power was modulated to a greater extent by the Self-produced finger-taps, as compared to Other-produced taps. In contrast with sustained beta suppression related to upregulation of primary motor cortex (Miller et al., 2007; Pfurtscheller & Lopes da Silva, 1999), frequency-specific phasic modulation is thought to reflect activity of sensorimotor networks dedicated to top-down control, prediction, and integration (M. Seeber et al., 2016). We found that coupling led to a significant general increase in beta modulation, as indicated by the main effect of Coupling in our statistical model. However, it should be noted that the interaction effect, while not statistically significant, strongly suggests a greater increase for Other as compared to Self (see *Figure 5.4B*).

We propose that the general increase in modulation is due to the greater task demand for attentional resources, which are re-allocated to predict the movements of the perceived partner. In previous work(Rosso, Maes, et al., 2021), we have shown that the exposure to the partner's competing rhythm systematically resulted in stable states of collective behavior, at the expenses of the individual synchronization performance. This is in line with the participants' self-reports, who rated coupled conditions as significantly more difficult compared to the uncoupled counterparts. Interestingly, participants also reported a significantly reduced sense of control over their own actions. Taken together, the evidence supports that beta modulation is more than a mere perceptual effect, and may reflect how the sensorimotor system is actively tracking perceived movement during a rhythmic interaction.

Our results fit well with the idea that ~20 Hz oscillations fulfill the function of attending rhythms, and that their power modulation reflects a dynamic allocation of neural resources (Nobre & van Ede, 2018) distributed among production (Self) and prediction (Other) of rhythmic movements. The mere perception of participants' taps required an additional deployment of attentional resources to comply with the task at hand, which could explain the increase of beta modulation by Self in condition of visual coupling, while the same component was significantly modulated by the unattended Other. Human rhythmic movement is known to convey exceptional salience, and as a privileged class of motor affordance (Phillips-Silver et al., 2010) it could compete with self-produced behavior for driving the attending rhythms. Furthermore, overt attention is not a necessary condition for the brain to perform efficient predictions, as beta modulation is partially independent from the attended sensory stream of information (van Ede et al., 2010). Our results show that the phenomenon is supramodal, as it occurred regardless of whether the metronome and the partner were visually or auditorily perceived. This extends recent evidence from a multimodal finger-tapping task (Nijhuis et al., 2021) to the interpersonal domain, further supporting that beta dynamics are mainly driven by top-down motor processes.

The observation that mutual modulation does not occur in the alpha band supports the frequency-specificity of the phenomenon for beta, with relevant functional implications. Despite the importance traditionally attributed to alpha dynamics for interpersonal coordination among coupled individuals (Tognoli & Kelso, 2015; Tognoli et al., 2007), our results suggest that a 'rhythmic mode' of attention is at play, rather than a general attentional or inhibitory mechanism. Our findings are in line with previous evidence for oscillatory dynamics in the beta range, which under certain experimental conditions are shown to encode the temporal regularities of predictable stimuli and enable predictive synchronization behavior. In light of the rhythmic nature of our task, we argue that the

neural encoding of movement periodicities in the beta dynamics, which in turn enables mutual predictions, represents a parsimonious interpretation of our findings. On a cautionary note, we want to state that our experiment was not designed to either prove or disprove any general theory of brain functioning: our occasional references to predictive coding and dynamic attending theory are only meant to present alternative influential perspectives in the state of the art.

The present work could not provide reliable information at the level of brain sources, due to the lack of co-registration of structural scans of the participants. However, the observation of systematic topographical features at the sensor level can guide some inferences on different brain networks involved across conditions. *Figure 5.5A* shows the grand-average spatial activation patterns for the ~ 20 Hz beta component as extracted via eigendecomposition. Focusing on the systematic differences and similarities across conditions, it is evident that the bilateral temporo-parietal activation systematically appears as the common pattern. In relation to the task, the only constant is the performance of finger-tapping paced by a rhythmic cue, whereas the nature of the cue and the coupling with the partner vary within the experimental design. However, when looking at the coupled conditions, an additional lateralized cluster emerges over left parietal regions. The subtraction of the activation coefficients (see *Figure 5.5B*) shows that the additional cluster is common across visual and auditory coupling modalities, pointing to the involvement of a supramodal network specific to the dyadic interaction.

Without drawing conclusions on the specific brain structures involved, we want to highlight that the additional cluster suggests the involvement of a network selective to the interaction, on top of the one dedicated to the finger-tapping task. This is in line with the idea of brain structures selectively involved in predicting human behavior in social contexts (Chris D. Frith, 2007), here activated in coupled conditions. Future work should be carried out to provide conclusive results on the actual underlying brain networks, possibly taking advantage of magneto-encephalography and combining structural magnetic resonance imaging co-registration with EEG source localization algorithms (Chella et al., 2019).

Our study is a novel investigation of coupling the brain activity of interacting participants to their reciprocal rhythmic behaviors. This approach uncovered a frequency-specific dynamic in the narrow frequency range around 20 Hz, which underpins individual production and mutual predictions of finger-tapping. By deploying dual-EEG recordings in an explicitly embodied approach to social neuroscience (Hamilton, 2021), we provided promising evidence for a common neural mechanism enabling interpersonal coordination.

Our focus on oscillatory dynamics should be integrated by future investigation at the level of brain networks, backed by complementary neuroimaging techniques.

Author Contributions

Mattia Rosso: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing – Original Draft, Visualization, Project administration, Funding acquisition. **Ole Adrian Heggli**: Conceptualization, Software, Validation, Resources, Writing – Review and Editing, Visualization. **Pieter-Jan Maes**: Writing – Review and Editing, Supervision. **Peter Vuust**: Conceptualization, Writing – Review and Editing, Supervision, Project administration. **Marc Leman**: Writing – Review and Editing, Supervision, Project administration.

Acknowledgments

The present study was funded by Bijzonder Onderzoeksfonds (BOF) from Ghent University (Belgium). The hardware of the recording device was built by Ivan Schepers. *Figure 5.1* and *5.2* were illustrated by Kevin Smink.

6

Neural entrainment meets behavior: stability index as a neural outcome measure of auditory-motor coupling

Mattia Rosso [A,B], Marc Leman [A], Lousin Moumdjian [A,C,D]

[A.] IPEM - Institute for Systematic Musicology; Ghent University; Ghent, Flanders, 9000; Belgium.

[B.] PSITEC - Psychologie: Interactions, Temps, Emotions, Cognition - ULR 4072; University of Lille; Lille, Hauts-de-France, 59650; France.

[C.] UMSC Hasselt-Pelt, Belgium

[D.] REVAL Rehabilitation Research Center, Faculty of Rehabilitation Sciences, Hasselt University

Rosso, M., Leman, M., & Moumdjian, L. (2021). Neural entrainment meets behavior: the stability index as a neural outcome measure of auditory-motor coupling. Frontiers in Human Neuroscience, 15.

Introduction

Auditory stimuli such as music or metronomes can entrain human movement, and this phenomenon can be used for neurological rehabilitation purposes. Particularly, evidence has been established that auditory stimuli can facilitate walking in persons with Parkinson's Disease (Ghai, Ghai, Schmitz, & Effenberg, 2018), stroke (Yoo & Kim, 2016) and multiple sclerosis (Moumdjian, Moens, Maes, Van Nieuwenhoven, et al., 2019). Auditory stimuli convey temporal structures that serve as affordances for the motor system to interact with (Marc Leman, 2016). In our previous work, we showed that auditory rhythms can entrain a person's motor rhythms, thus affecting abilities for walking. The underlying mechanism can be explained in terms of sensorimotor phase-locking, prediction error minimization, and/or dynamical interactions (Marc Leman, 2016; Phillips-Silver et al., 2010). The outcome of an entrainment process is typically a more stable state of synchronization (Marc Leman, 2016; Bart Moens et al., 2014). So far, the entrainment effect has been quantified by means of behavioral outcome measures, in particular temporal outcomes of the rhythmic auditory-motor coupling (Moumdjian, Buhmann, Willems, Feys, & Leman, 2018), which contributed to a better understanding of underlying mechanisms as a result of the interaction (Moumdjian et al., 2020; Moumdjian, Moens, Vanzeir, et al., 2019), and to the development of task-oriented training tools for walking in persons with the neurological disease of multiple sclerosis (Moumdjian, Moens, Maes, Van Geel, et al., 2019; Moumdjian, Moens, Maes, Van Nieuwenhoven, et al., 2019).

Part of the variability in entrainment can be attributed to individual synchronization abilities. When presented with auditory stimuli and asked to walk to them, there are participants who spontaneously synchronize, and others who do not. This ability is not only limited to neurological populations, but also holds true for healthy participants (Van Dyck et al., 2015) where the percentage of spontaneous synchronizers versus non-synchronizers is about 50%-50%. A number of factors contribute to the tendency to rhythmically entrain and synchronize (Wilson & Cook, 2016). The first factor is temporal perception and prediction. Studies on Parkinson's Disease have concluded that those participants with higher perceptual sensorimotor synchronization abilities, quantified by behavioral sensorimotor tapping tasks involving finger tapping (Dalla Bella et al., 2017), had a better outcome on their walking parameters after being subjected to walk to auditory stimuli (Bella et al., 2017). The second factor is motor (e.g., physical capacity) and/or cognitive (e.g. attentive and pre-attentive) functions. For example, studies comparing spontaneous and instructed synchronization of walking (Leow, Waclawik, & Grahn, 2018; Moumdjian, Moens, Vanzeir, et al., 2019) and running (Van Dyck, Buhmann, & Lorenzoni, 2021) to music have been conclusive that explicit instructions to synchronize

resulted in a higher synchronization tendency, as compared to spontaneous synchronization. The former study also noted this difference across different motor thresholds which was provided as a result of walking to different tempi, starting from the natural comfortable tempo and up to +10%, in increments of 2% (Moumdjian, Moens, Maes, Van Nieuwenhoven, et al., 2019; Moumdjian, Moens, Vanzeir, et al., 2019).

Up to now, most studies on neurological populations, investigating entrainment and synchronization during walking tasks, are based on empirical evidence using behavioral outcomes (Moumdjian et al., 2018). However, we believe that the development of complementary neurological outcomes could offer a further understanding of entrainment and synchronization, potentially leading to the development of more individualized and more fine-tuned rehabilitation approaches.

The present study therefore aims at quantifying a neural outcome measure of entrainment and synchronization in combination with behavioral outcomes. We propose the use of electroencephalography (EEG) as a method to measure neural entrainment of the motor system to rhythmic stimuli. The novel outcome measure is based on a finger tapping task (M. L. Bavassi, Tagliazucchi, & Laje, 2013; López & Laje, 2019; McPherson, Berger, Alagapan, & Fröhlich, 2018). *Figure 6.1* shows a graphical illustration of this study's rationale and proposed contribution to the current state of the art.

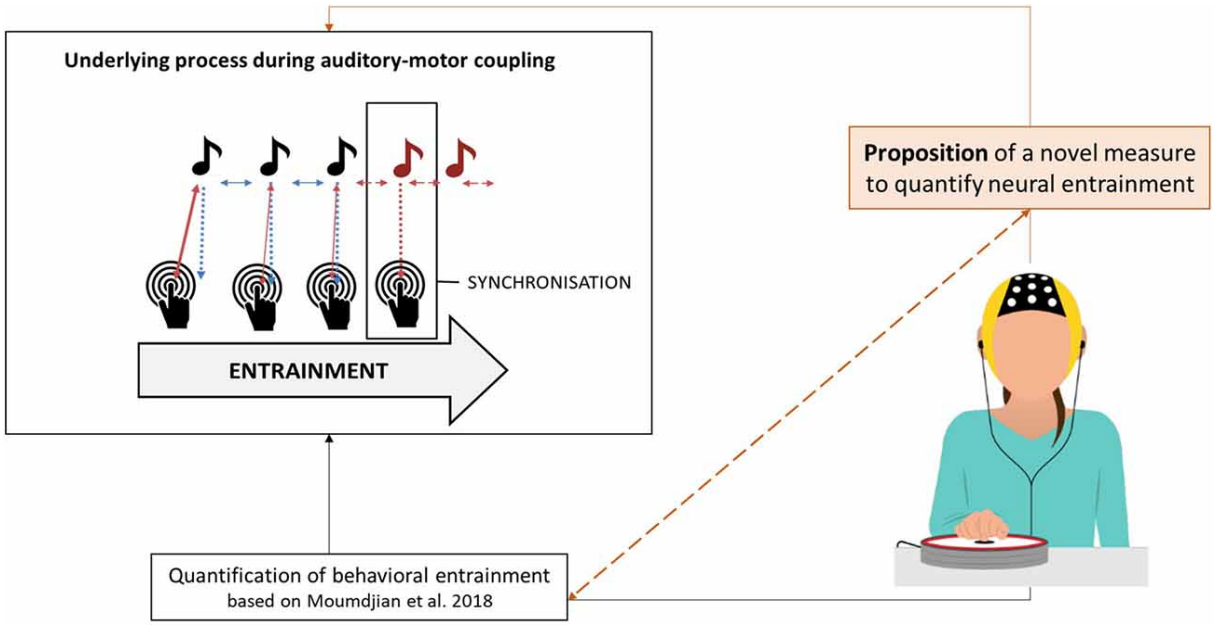


Figure 6. 1. Rationale of the study and contribution to the state of the art.

Our approach is based on Steady-State Evoked Potentials (SSEPs) (Norcia et al., 2015; Vialatte et al., 2010). Given a steady periodic stimulation, a series of subsequent evoked responses is elicited in the electrical brain activity, generating a periodic pattern of transients in the EEG signal. By transforming the signal to the frequency domain by means of Fast Fourier Transform (FFT), it can be observed that the EEG spectrum is dominated by a prominent peak at the stimulation frequency and its harmonics. Upon exposure to rhythmic auditory stimuli, patterns emerge in brain activity and match the dominant spectral features of the stimulation. Studies show that neural entrainment can be measured at different hierarchical levels of the stimulus temporal structure, or of its representation (Nozaradan et al., 2011). As the sound envelope of a musical stimulus exhibits a periodic low-frequency amplitude modulation in correspondence with the beat, it is possible to observe a match between the beat-related harmonics of the EEG spectrum and the sound spectrum (Lenc, Keller, Varlet, & Nozaradan, 2018). However, the observed entrained components are not always entirely driven by sensory stimulation. In fact, given the same energy in the stimulus, SSEP amplitude is modulated by attention (S. K. Andersen, Fuchs, & Müller, 2011; Kashiwase, Matsumiya, Kuriki, & Shioiri, 2012), internal representation of metric structure (Nozaradan, Peretz, & Mouraux, 2012), sensorimotor integration (Nozaradan, Zerouali, Peretz, & Mouraux, 2015) and interpersonal coordination (Varlet et al., 2020).

The SSEP technique is relatively straightforward in modelling bottom-up and top-down components of rhythm perception in terms of Fourier coefficients. However, in order to link behavioral entrainment to a neural outcome measure, we believe that the signal phase should not be left out of the picture. Rajendran and Schnupp (Rajendran & Schnupp, 2019) showed that shuffling the phase of a signal resulted in drastic differences in its time domain representation, whereas it remained invariant in the frequency domain. Although the analysis of peak amplitudes or z-scores in a static spectrum might convey information about the outcome of neural entrainment, it is arguably insensitive to its dynamics in the time domain. One should consider that oscillatory processes in the brain are hardly stationary (Michael X. Cohen, 2017) and the very definition of entrainment implies that an oscillator dynamically changes its frequency in order to achieve stable synchronization. This is precisely the phenomenon we intend to capture. Therefore, in order to quantify neural entrainment of rhythmic stimuli, we argue in favor of a time-varying measure based on the phase of the neural entrained component.

With this study, we progress beyond the state of the art in the research on neural entrainment by optimizing the calculation of a neural outcome measure of auditory-motor coupling. We argue that such a measure can be used together with its behavioral counterparts. In combination, both the behavioral and neurological measures may unveil

a further layer of the underlying mechanisms of the rich dynamical processes during motor and auditory interactions. Our first aim is to extract from the EEG signal the component which is maximally entrained to a periodic stimulus. For that we compute a *stability index* to quantify frequency fluctuations over time. Our second aim is to validate the proposed index with a set of quantified behavioral outcome measures of auditory-motor coupling and entrainment (Moumdjian et al., 2018). In an auditory-motor coupling task, healthy participants were instructed to tap their index finger synchronizing to an auditory metronome (as illustrated in *Figure 6.1*). We hypothesized that our stability index would significantly correlate with the behavioral measures of entrainment. Specifically, a stable behavioral performance is expected to correlate with a stable entrained component, whereas a poor performance would result in wider frequency fluctuations over time.

Materials and Methods

Participants. Twenty-eight (N = 28) right handed participants took part in the study (18 females, 10 males; mean age = 29.07 years, standard deviation = 5.73 years). None of them had a history of neurological, major medical or psychiatric disorders. All of them declared not to be professional musicians upon recruitment, although some of them had musical experience. Handedness was assessed by means of the Edinburgh Handedness Inventory (Oldfield, 1971). The experiment was approved by the Ethics Committee of Ghent University (Faculty of Arts and Philosophy) and informed written consent was obtained from each participant, who received a 15€ coupon as economic compensation for their participation.

Experimental procedure. The experimental task consisted of a tapping synchronization paradigm, in a sitting position. Participants were provided with a custom-made pad containing piezo sensors to detect tapping onsets, and were instructed to tap their right index finger along with the assigned metronome during 390 seconds. During the task, participants were sitting on a comfortable chair equipped with armrests, so that their elbow could lay in a fixed position. Tapping movements were limited to wrist flexion in order to prevent movement-artifacts contamination of the EEG signal. Participants were monitored on-line and video-recorded by means of a USB-camera to verify their compliance with the instructions. The importance of avoiding head and trunk movements was stressed.

***Auditory stimuli*.** Participants were presented with the stimuli via DefenderShield® air-tube ear plugs. Ableton Live 10 ® was adopted as software for the metronome stimuli presentation. A periodic auditory cue was presented at a rate of 100 BPM to half of the participants, and 98.5 BPM to the other half (1.67Hz and 1.64Hz respectively). The reason of such minimal gap lies in the rationale of a larger experimental design in which the recordings were performed (Rosso, Maes, et al., 2021).

Behavioral data acquisition. Finger tapping onsets were recorded with a Teensy 3.2 microcontroller, operating as serial/MIDI hub in the setting. On the one hand, it received an analogic input from piezo sensors inside the pads and printed on the serial port of the stimulation computer a timestamp each time a finger-tap pushed the signal above a resting threshold. The threshold was conservative enough to prevent false positives due to signal bouncing. Every time a metronome beat onset was presented to a participant, a MIDI message was sent to the Teensy to log its timestamp on the serial port. All timestamps were rounded to 1ms resolution, which corresponds to 1kHz sampling rate. The same device triggered the start of the EEG recording by sending a TTL trigger via BNC connection.

Outcome Measures

Behavioral data and neurophysiological data were measured. These are outlined below.

Behavioral data. The timestamps of finger-tapping and metronome beat onsets were imported in Matlab® and used to calculate a set of behavioral outcome measures of auditory-motor coupling and entrainment. Before doing so, we removed the finger-tapping onsets following the previous one by less than 350ms, as false positives could occasionally be recorded when a participant pushed the pad for too long or accidentally laid the hand on it. On average, 0.4 false positives were removed for every participant (standard deviation = 0.8). From the finger-tapping and metronome beat onsets time-series, we calculated the following measures: relative phase angle, resultant vector length, mean asynchrony and tempo matching. Below, details of the measures and the formulae used calculate these measures are outlined (Bart Moens et al., 2014; Moumdjian et al., 2018):

Relative phase angle. This is an error measure of synchronization based on the phase difference between two oscillators (i.e., the participant tapping and the metronome beat onsets).

$$\varphi = 360 * \left(\frac{T_n - B_n}{B_{(n+1)} - B_n} \right)$$

Where T_n is the participant's tap onset *n* and B_n is the onset of the closest metronome beat. A negative angle indicates that the participant is tapping ahead of the metronome, while a positive angle indicates that the participant's tap is lagging behind. Alternatively, following recent work on modelling participants and periodic cues of systems of coupled oscillators in finger-tapping studies (Ole Adrian Heggli et al., 2019), we processed the phase time-series for participants and metronomes by interpolating the onsets as a ramp wave, wrapped from 0 to 2π radians at 1kHz sampling rate. Provided with an estimate of the oscillators' positions on their cycle with a temporal resolution of 1ms, we subtracted each participant's phase time-series from the respective metronome. Finally, the *CircStats* toolbox[34] for Matlab® was used to calculate the mean angle from the resulting relative phase time-series (in radians).

Resultant vector length. This expresses the stability of the relative phase angles over time. A unimodal distribution implies a high resultant vector length, whereas uniform and bipolar distributions result in a low resultant vector length. The measure was processed with the *CircStats* toolbox (Berens, 2009), using the relative phase time-series as input. The measure ranges from 0 to 1, where 1 indicates perfect synchronization over time at a given relative phase angle, and is calculated as follows:

$$R = \left| \frac{1}{N} \sum_{n=1}^{N} e^{i\phi T_n} \right|$$

Mean asynchrony. This consists of the mean difference between the participant's tap onsets and the respective closest metronome's beat onset expressed in milliseconds.

$$Mean\ asynchrony = \frac{1}{N}\sum_{n=1}^{N} T_n - B_n$$

Tempo matching accuracy. This indicates to what extent the overall tempo of the participant's tapping matches with the tempo of the metronome beats, based on inter-onset-intervals (IOIs). Inter-beat deviation (IBD) is calculated as the standard deviation of a subject's IOIs with respect to the metronome.

$$IBD = \frac{1}{N}\sum_{n=2}^{N} \frac{\left(B_n - B_{(n-1)}\right) - \left(T_n - T_{(n-1)}\right)}{B_n - B_{(n-1)}}$$

Neurophysiological data. In order to compute our proposed outcome measure of neural entrainment, the 'stability index', the following EEG processing pipeline was conducted. It consists of signal pre-processing, generalized eigendecomposition (GED) and computation of the stability index. The steps leading to the computation of our final measure are summarized in *Figure 6.2*:

Data acquisition. Participants were equipped with a 64-channel *waveguard™original* EEG headset (10-10 system, with Ag/AgCl electrodes). Data were recorded with an ANT-Neuro *eego™mylab* system at 1kHz sampling rate. Impedances were monitored in the *eego™* software environment and kept below 20kΩ. In comparison with stricter thresholds (e.g., 5kΩ or 10kΩ), the choice made it feasible to maximize the homogeneity of impedance levels across electrodes, and in turn optimize the covariance matrices used in our source separation. A referential montage was adopted, with 'CPz' as the reference electrode.

Pre-processing. Pre-processing was carried out with a pipeline integrating functions from the *Fieldtrip toolbox* (Oostenveld, Fries, Maris, & Schoffelen, 2011) for Matlab (MathWorks Inc, USA). Bad channels were identified by means of visual inspection of raw

timeseries and variance distribution across channels. The recordings were re-referenced to the average of all the electrodes after channel rejection, to avoid noise leakage into the average. A high-pass Butterworth filter with 1Hz cut-off was applied to the raw recordings to remove slow drifts. We preferred to choose this conservative threshold, given that occasional head movements and sweat potentials are more likely to occur over a long continuous recording. A low-pass Butterworth filter with 45Hz cut-off was applied to remove high-frequency muscular activity. A notch filter centered at 50Hz was applied to remove power-line noise up to the 3rd harmonic.

Independent component analysis (ICA) was conducted on full rank data to remove blinks and eye-movement artifacts, by means of visual inspection of topographical maps and time-series of component activation. For this purpose, we run the 'runica' algorithm as implemented in Fieldtrip, excluding the reference 'CPz' and the bad channels timeseries from the input matrix. Only those components which exhibited the stereotyped frontal distribution generated by blinks and lateral eye movements were removed. Although other artifactual sources could have been identified, we limited the selection to a few unambiguous components for the sake of replicability. A minimum of 1 and a maximum of 3 components were removed for every participant. The dataset was inspected prior to ICA decomposition and following ICA back-projection. Special attention was given to the electrodes where the activation of the artifactual component was maximal, namely the F, AF, and Fp clusters. Rejected bad channels were finally reconstructed after artifact removal, by computing a weighted average of all neighbors as implemented in Fieldtrip.

Recordings were treated as a continuous experimental run, without segmentation in epochs. This implies that no 'bad trials' were removed. Further in this section, we will present how we dealt with transient bursts of artifactual activity in the continuous recording.

Generalized Eigendecomposition (GED). In order to avoid channel selection bias while optimizing the signal-to-noise ratio between the entrained component and the broadband neural activity, we applied GED as first described in the context of source separation for rhythmic entrainment (Michael X. Cohen & Gulbinaite, 2017). The technique consists of a spatial filter to reduce the multivariate dataset to one dimension, guided by some criteria: in this case, it was attunement to the stimulation frequency. This is achieved by computing the weighted average of a set of channels, where the vector of weights *W* was calculated by solving the following eigenequation:

$$R^{-1}SW = \Lambda W$$

where *S* is the covariance matrix calculated from the narrow-band filtered signal; *R* is the reference covariance matrix calculated from the broad-band signal; Λ is a set of eigenvalues. GED identifies eigenvectors W that best separate the signal ('S') covariance from the reference ('R') covariance matrix. The eigenvector associated with the largest eigenvalue was selected as spatial filter. That eigenvector is then used to multiply the raw channel data to produce the single time series of our target entrained component. In the present work, a subset of 37 channels located behind the frontocentral 'FC' line (mastoids excluded) was selected. By doing so, we intended to constrain the source separation and target a sensorimotor component entrained to the auditory stimulus. The excluded channels form the cluster which is typically expected from a purely auditory response at the scalp level (Nozaradan et al., 2011, 2015). The regions of interest (ROIs) selection is visually illustrated in *Figure 6.2*.

Given we were explicitly looking for frequency fluctuations, our narrow-band filter needed to be large enough in order to leave room for fluctuations around the entrained frequency. We designed our filter as a Gaussian function in the frequency domain, with center at 1.654 Hz and a width of 0.3 Hz at half of the maximum gain. Such parameter represents an optimal trade-off in our application, since it allows for fluctuations around the center frequency without overlapping with the high-pass band filter (cut-off = 1 Hz). We then filtered the signal on the whole subset of 37 channels by performing element-wise multiplication between the signal spectrum and the filter kernel. The resulting spectrum was eventually transformed with inverse Fast Fourier Transform back in the time domain. The frequency-domain representation of the filter kernel is provided in *Supplementary Figure 6.A*, along with additional information in the figure's description.

The reference R covariance matrix was here computed from the broadband multivariate signal. Our choice differs from the approach originally proposed by Cohen and Gulbinaite (Michael X. Cohen & Gulbinaite, 2017) in that they propose a use-case for higher frequency ranges, which allows to average the R matrices computed from two narrow-band Gaussian flankers neighboring the central filter on both sides. Their rationale is to minimize the contribution of intrinsic non-task related rhythms in frequency ranges far from the one of interest, while avoiding bias from upper and lower frequency neighbors. Given that we were dealing with low frequencies (< 2 Hz), it was not desirable for us to narrow-band filter the signal in a lower flanker, as we would have reached below the high-pass filter cut-off at 1 Hz (see *Supplementary Figure 6.A*). In case we adopted flankers, we would have had a bias to the right side of the spectrum.

In order to compute the respective covariance matrices from the broad- and narrow-band signals, we used the onset timing of the finger-taps performed by the participant to define time-windows from -100ms to 500ms around the events. The approach provided us with a considerable number of covariance matrices for each recording (645 finger-taps were expected on average), such that we could remove the ones whose Euclidean distance from the grand-average covariance matrix exceeded the 2.23 z-scores (i.e., corresponding to a probability of 0.013). The grand-average S and R covariance matrices where then calculated free from the occasional burst of artifactual activity over the long recording, compensating from the impossibility of performing a procedure of trial-removal during the preprocessing. The quality of our GED application was assessed by inspecting the eigenspectrum, the topographical activation map and the power spectrum of the extracted component (see *Figure 6.3*).

Stability Index. Once the entrained component was computed, we applied on it the same Gaussian filter (center at 1.654 Hz and 0.3 Hz width at half maximum) in order to extract reliable phase time-series from the analytic signal. We calculated the analytic signal with the Hilbert transform and computed the instantaneous frequency time-series from the first derivative of the unwrapped phase angles time-series as indicated in (Michael X. Cohen, 2014). The instantaneous frequency of a dynamical oscillating system can be defined as the change in the phase per unit time (Boashash, 1992). The derivative can then be converted to Hz applying the following formula:

$$Hz_t = \frac{s(\phi_t - \phi_{t-1})}{2\pi}$$

where s indicates the data sampling rate and ϕ_t indicates the (unwrapped) phase angle at time *t.* A sliding moving median with a window width of 400 samples was used to smooth the instantaneous frequency time-series, to remove occasional extreme bursts due to artifactual activity distorting phase time-series. Finally, we calculated the standard deviation of instantaneous frequency over the whole task as a global measure of frequency stability over time, which we named the '*stability index*'. A high standard deviation is thus indicative of wide instantaneous frequency fluctuations, and less overall stability of the entrained component. A standard deviation equal to 0 indicates a perfectly stable component, with the instantaneous frequency being a flat line at the constant value of the stimulus frequency.

Statistical analysis. In order to validate our neural outcome measure, we calculated the Spearman coefficient for the correlation between the stability index and the four behavioral outcome measures reported above. This technique assesses the strength and significance of monotonic relationships between variables, regardless of its linearity. The Spearman correlation coefficient computed on continuous variables is the equivalent of the Pearson correlation coefficient computed on their ranks: it is exempt from the assumption of normal distribution of the pair of variables, and robust to outliers and scaling effects. The following classification was used to categorize the correlation (Witz, Hinkle, Wiersma, & Jurs, 1990): .00-.30 'negligible correlation', .30-.50 'low correlation', .50-.70 'moderate correlation', .70-.90 'high correlation', .90-.100 'very high correlation'.

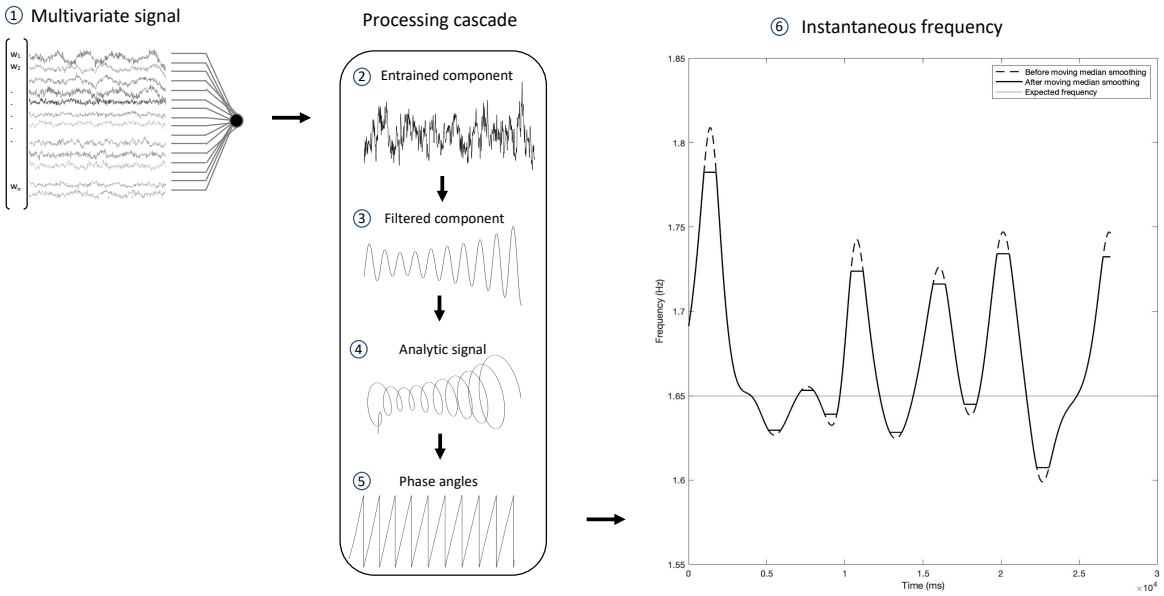


Figure 6. 2. EEG processing pipeline. The present pipeline illustrates the steps through which the proposed stability index was computed. Following the pre-processing, the vector of weights w associated with the highest eigenvalue was used as a spatial filter: by multiplying the data from the 35 channels behind the frontocentral line (1), we produced a single time series. The resulting *entrained component* (2) went through a cascade of computational steps: first, it was narrow-band filtered with a Gaussian filter centered at the stimulus frequency, in order to extract reliable phase timeseries unaffected by broad-band component (center = 1.65 Hz; width at half-maximum = 0.3 Hz). The *filtered component* (3) was then Hilbert-transformed to produce the *analytic signal* (4), from which we computed the *phase angles* timeseries (5). Finally, the phase was unwrapped, its first derivative was used to compute the *instantaneous frequency* (6), and a sliding moving median was applied in order to level out eventual artifactual peaks. The plot shows how the pipeline results in a time-varying measure of frequency over time, which fluctuates around the stimulation frequency (i.e., the thin horizontal line intercepting the y-axis at 1.65 Hz). The standard deviation of the instantaneous frequency provides a global measure of stability of the entrained component for a given time window, which in our case was the whole duration of the task. We named such global measure *stability index*, for it equals 0 in case of a flat horizontal line. Such scenario would be observed in the ideal case of a perfectly stable component oscillating like a simple sinewave.

Results

Behavioral outcome measures

On a group level, we report that participants anticipated their tapping onsets relative to the beat, with a mean relative phase angle of -1.050 ± 0.681 radians and a mean asynchrony of -77.472 ± 40.603 milliseconds. In addition, on a group level, they obtained a consistent synchronization with a relative vector length of 0.831 ± 0.156, and a consistent period measured by the inter-beat deviation of – 0.001 ± 0.01. The individual participant behavioral results of these outcomes are reported in *Table 6.1*.

Neural outcome measure

Generalized Eigendecomposition. The source separation successfully extracted the entrained neural component of interest, as assessed by its spectral features and its topographical map of activation. The component associated to the higher eigenvalue was selected for our analyses. Additionally, we verified that the component associated with the second eigenvalue was not related to the behavioral performance. More details about the second component are provided in the *Supplementary material 6.B* and *6.C*. A detailed profile of the first component is provided in *Figure 6.3*, and its functional meaning will be further discussed in the next section.

Stability Index. The stability index was computed as the standard deviation of the component's instantaneous frequency, as described in the Methods section. On a group level, the stability index resulted in 0.062 ± 0.030 Hz. A stability index of 0 indicates a perfectly stable component, without any frequency fluctuation over time. The individual participant results of the stability index are reported in *Table 6.1*.

Correlation analysis

As shown in *Figure 6.4*, the Spearman correlation between the behavioral outcome measures of entrainment and the stability index revealed significant moderate negative correlations for relative phase angle, resultant vector length and mean asynchrony ($r = -0.566$, $p < .001$; $r = -0.652$, $p < .001$; $r = -0.523$ $p = .005$ respectively). A non-significant negligible correlation was found for the inter-beat deviation ($r = 0.107$, $p = .583$).

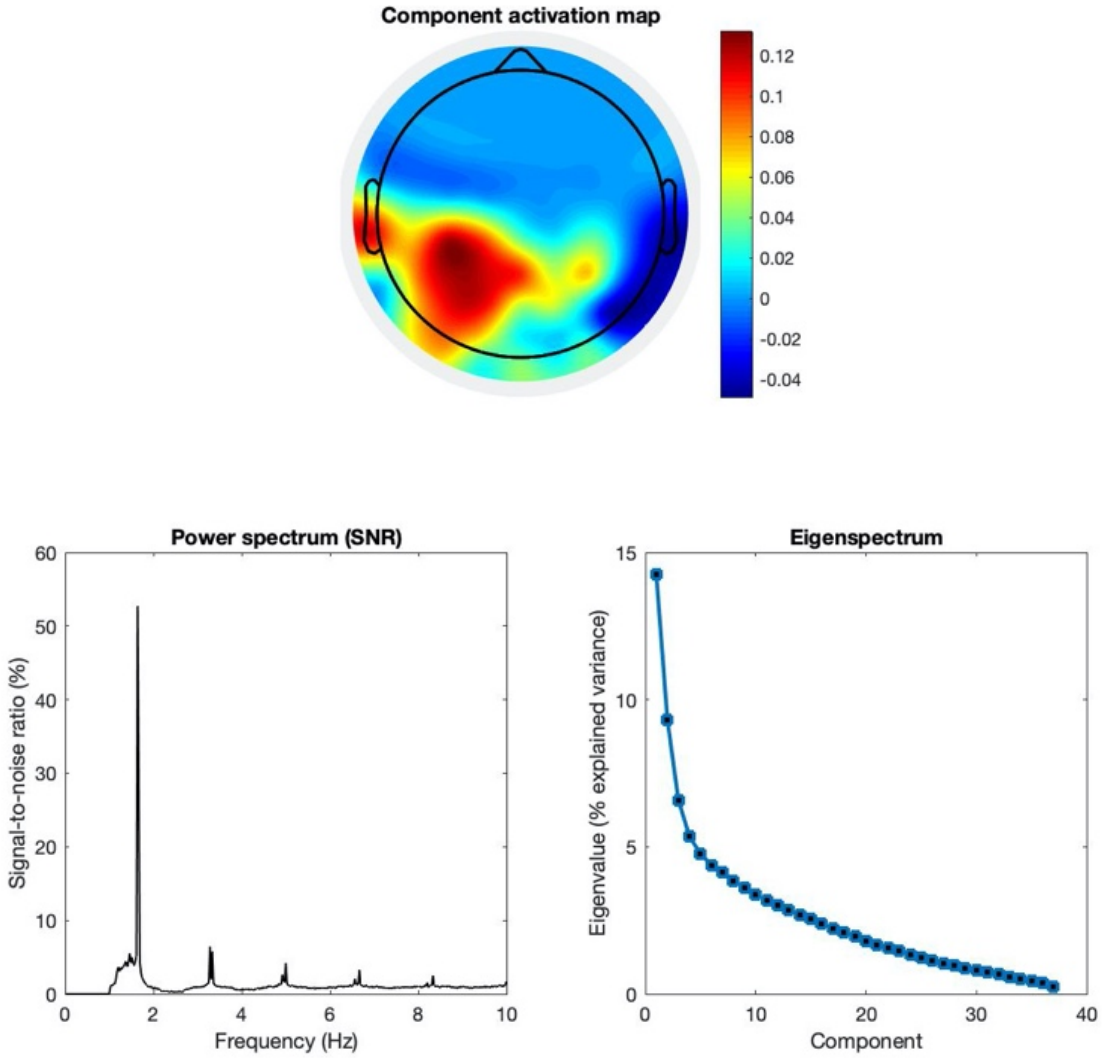


Figure 6. 3. Group-level assessment of the source separation. The following criteria were used to assess the quality of our source separation via GED: A) *Topography*. The grand-average coefficients of activation are shown in the topographic plot: maximal activation was recorded at the left centroparietal 'CP' cluster and at left temporal electrodes ('T7' and 'TP7'). It should be noted that we explicitly excluded from the spatial filter the channels located beyond the frontocentral line, for we intended to maximize an entrained response related to sensorimotor processing in the context of the task. B) *SNR spectrum*. The grand-average power spectrum is here represented in percentage signal-to-noise ratio between each data point and the mean power in the flanker bins (0.5 Hz), in order to remove the physiological 1/f component of the spectrum (Freeman, Holmes, Burke, & Vanhatalo, 2003). C) *Eigenspectrum*. The grand-average eigenvalues are sorted in descending order, showing an exponential decay with the vector of weights w used as spatial filter is the one associated with the highest eigenvalue λ. Before averaging, eigenvalues where normalized and expressed as percentage of explained variance. All grand-averages were computed on the whole sample of participants (N = 28).

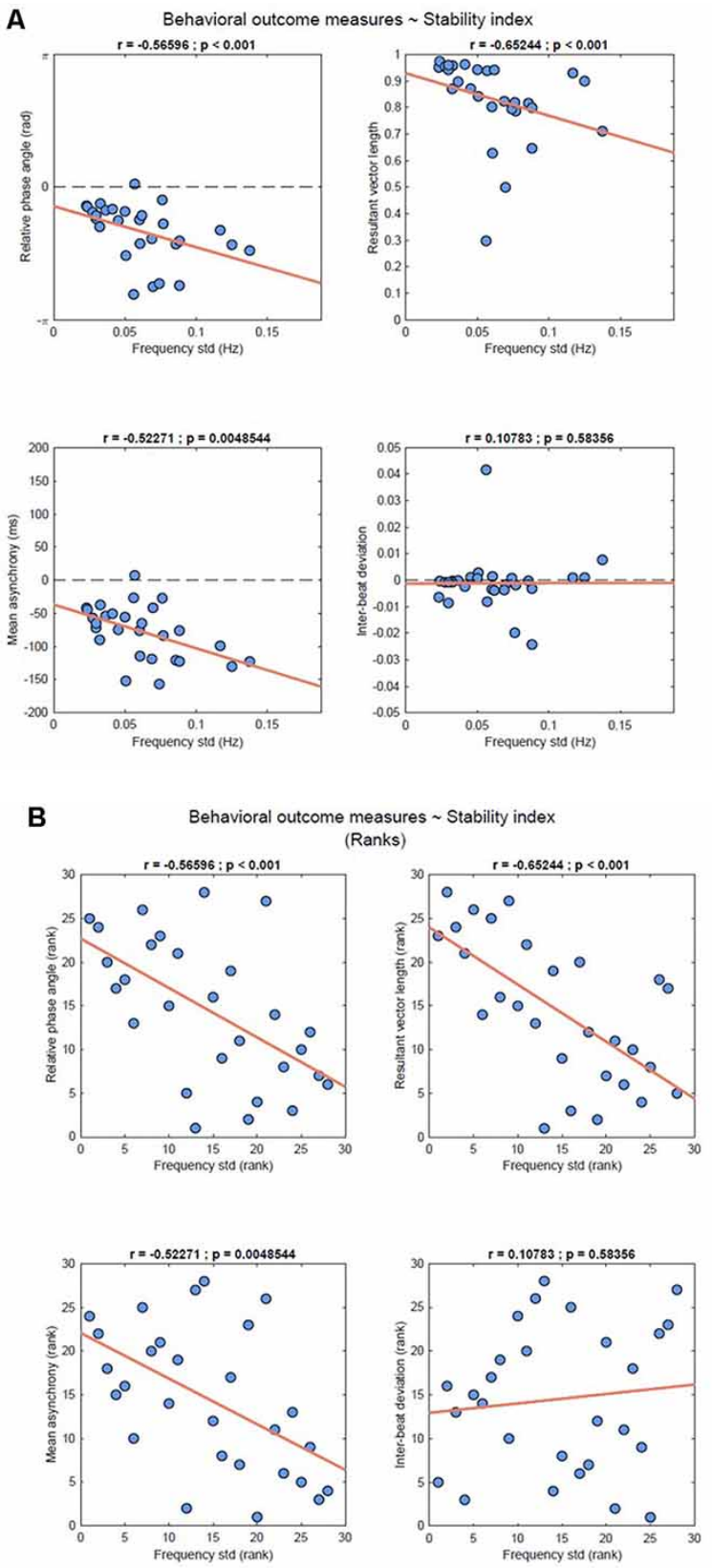


Figure 6. 4. Brain-behavior correlations. A) Results of the Spearman's correlation analysis between the behavioral outcome measures and stability index, for all study participants. Data are represented on the original scale. B) Correlations between the ranks for the behavioral outcome measures and stability index of all study participants.

Participant ID	Neural outcome measure of entrainment	Behavioural outcome measures of entrainment			
	Stability Index (frequency fluctuation - std)	Relative phase angle (radians)	Resultant vector length (0-1)	Mean asynchrony (ms)	Inter-beat deviation (ratio)
1	0.033	-0.387	0.958	-37.259	0.000
2	0.088	-2.332	0.647	-76.046	-0.003
3	0.023	-0.436	0.952	-41.668	-0.006
4	0.024	-0.464	0.975	-44.403	0.000
5	0.057	0.078	0.939	7.412	-0.008
6	0.138	-1.503	0.711	-122.858	0.008
7	0.077	-0.867	0.787	-83.514	-0.002
8	0.060	-0.772	0.803	-76.068	-0.003
9	0.056	-2.543	0.298	-26.583	0.042
10	0.061	-1.337	0.629	-114.613	0.001
11	0.037	-0.547	0.898	-53.906	0.000
12	0.086	-1.349	0.817	-120.805	0.000
13	0.070	-2.358	0.500	-41.690	-0.002
14	0.088	-1.273	0.800	-122.573	-0.024
15	0.045	-0.790	0.872	-74.723	0.001
16	0.125	-1.364	0.900	-130.310	0.001
17	0.051	-1.621	0.843	-152.019	0.003
18	0.028	-0.590	0.955	-57.088	-0.001
19	0.030	-0.744	0.943	-72.028	-0.009
20	0.074	-2.285	0.795	-156.864	0.001
21	0.076	-0.301	0.820	-27.093	-0.020
22	0.050	-0.574	0.943	-55.641	0.001

23	0.069	-1.223	0.824	-118.943	-0.004
24	0.032	-0.937	0.871	-90.031	-0.001
25	0.030	-0.674	0.960	-65.188	-0.001
26	0.041	-0.519	0.963	-50.530	-0.002
27	0.117	-1.018	0.931	-99.138	0.001
28	0.062	-0.670	0.942	-65.033	-0.004

Table 6. 1. Results of neural and behavioral outcome measures of entrainment, per participant.

Discussion

The main contribution of the present work is methodological, motivated by the need to compute a neural outcome measure of neural entrainment in the context of auditory-motor coupling and prospectively applying auditory-motor coupling paradigms for the purpose of neurological rehabilitation. We proposed a novel processing pipeline to compute the *stability index*, and validated this neural outcome measure by testing its correlation with a set of behavioral outcome measures in the context of a finger-tapping task.

Behaviorally, participants exhibited the mean negative asynchrony typically reported in finger-tapping synchronization tasks performed by healthy participants (Gisa Aschersleben, 2002). The mean negative asynchrony and the negative relative phase angles confirmed that all participants but one tapped on average ahead of the metronome, anticipating the beat. Additionally, by looking at the resultant vector lengths, we also note that consistent synchronization was maintained throughout the task. Given these results, we can deduce that all subjects were engaged in the process of entraining their finger-taps to the auditory beats of the metronome.

As for the data captured by the EEG, our GED implementation was effective in extracting the target component maximally entrained to the rhythmic stimulus. *Figure 6.3* provides a quality-check for our source separation by combining the following three criteria at the group-level. The first, *topography*: the grand-average activation map of the selected component shows maximal activity in the left centroparietal cluster and in the left temporal electrodes. Such distribution strongly suggests the involvement of primary sensorimotor areas, given it is contralateral to the effector (i.e. the right hand). The same pattern was previously reported for movement-related SSEPs in the context of overt synchronized behavior (Nozaradan et al., 2015), and clearly differs from the frontocentral topography typical of auditory cortical responses in absence of movement (Nozaradan et al., 2012). Given our focus on sensorimotor dynamics underlying overt behavior, our spatial filter was constrained within the whole set of channels located behind the frontocentral line. Secondly, the *power spectrum*: a single major peak stands out at the metronome's frequency, accompanied by harmonics whose power approximately follows a 1/f distribution. The dominance of the target frequency over the spectrum shows that the extracted component is effectively fine-tuned to the rhythmic stimulation. Thirdly, the *eigenspectrum*: by sorting the eigenvalues in descending order, it is evident how the first one eigenvalue stands out over the rest of the spectrum. Such condition is particularly desirable when the goal is to reduce the dimensionality of a multivariate dataset to one single component that satisfies a given criterion. The eigenvector associated with the

highest eigenvalue could then be reliably used to weight the electrodes average, and reduce the dimensionality of the dataset to one entrained component.

Applying GED in the context of neural entrainment is an established method of optimizing source separation (Michael X. Cohen & Gulbinaite, 2017), avoiding major drawbacks of electrode selection. To elaborate, we chose this approach instead of selecting a time-series based on single electrode or on a small cluster of electrodes in order avoid subjective judgment to some degree. Despite this drawback, electrode selection is a rather common practice in the SSEP literature (e.g., S. K. Andersen et al., 2011; S. K. Andersen, Müller, & Martinovic, 2012; Kashiwase et al., 2012; Keitel, Andersen, & Müller, 2010; Rossion, Prieto, Boremanse, Kuefner, & Van Belle, 2012). In addition, with our spatial filter, we A) decrease the risk of attenuating the response in some subjects due to individual variability, and B) are not confounded to exposure to noise which might selectively affect a single channel. Although it is true that computing a non-weighted average over the whole scalp is sometimes proposed as a practice to avoid selection bias (e.g., see Nozaradan, Peretz, & Keller, 2016), the entrained response would be heavily attenuated by broadband components unrelated to the task. On the other hand, a weighted average oriented by spectral criteria would clearly overcomes such issues. Most importantly, our methodology of GED application was optimal for single-trial analysis and provided us with a single time-series whose time-course and dynamics could be further analyzed. Such time-series represented the starting point of our pipeline towards the computation of the stability index (see *Figure 6.2*). It should also be noted that rhythmic motor acts such as finger-tapping (McAuley, 2010; Moelants, 2002) and walking (MacDougall & Moore, 2005) operate within the low delta frequency range (Morillon, Arnal, Schroeder, & Keitel, 2019), which implies that long trials are needed to measure the dynamics of slower oscillatory components.

In order to validate our neural outcome measure of auditory-motor coupling, we ran correlation analyses with a set of behavioral measures of synchronization accuracy and stability in the context of a finger-tapping task to a metronome's beats. The stability index exhibited moderate negative correlations with the relative phase angles, the mean asynchrony and the resultant vector length. To explain our results, we first provide an explanation of the pattern we observed in the context of the stability of the frequency fluctuations, which are used to quantify the stability index. We observed less stability in the frequency fluctuations of the neural entrained component when the finger-taps were further away and with a wider distribution relative to the beats as reported by the relative phase angles and resultant vector length, respectively. Conversely, when the finger taps were closer to and in anticipation of the beat, with a narrow distribution, we observed that the entrained neural component stabilized its frequency fluctuations. The results

confirmed the hypothesis that these frequency fluctuations, as quantified by the stability index, correlated with the behavioral outcome measures of entrainment.

With our results, we also observed that the stability index was selectively correlated with measures of phase error correction mechanisms, and not with those of period error correction. This is consistent with the fact that the stability index was not correlated to the inter-beat deviations, which is a measure for quantifying tempo matching (Moumdjian et al., 2018). In turn, tempo matching is an outcome which describes error correction in period. With the above explanations, our results are suggestive that the stability index quantifies neural entrainment, yet limited to corrections in phase. However, we do not rule out the possibility that we did not find any significant correlation due to the very low individual variability in inter-beat deviations, which resulted in a small slope of the regression line. The result indicates that participants were very accurate in matching the period of the metronome over the whole duration of the task.

The selectivity of these correlations further supports the relevance of temporal dynamics at the micro-timing scale. By picking up on the notion of "neural entrainment to the beat", which is traditionally inferred from the Fourier coefficients of a "static" spectrum, we developed it towards a phase-based measure to make it sensitive to the temporal structure of the stimulus (Rajendran & Schnupp, 2019) and to behavioral dynamics. From our standpoint, in order to entrain to the beat a neural component should not only be tuned to the stimulus frequency, but it should dynamically attune depending on the ongoing entrained behavior. The stability index proposed in this context shows how frequency fluctuates over time as a function of the distance from in-phase synchronization (the phase angles and asynchrony) and consistency of the established relative-phase during the course of the task (resultant vector length). Previous work provided evidence on correlation between cortical entrainment and overt sensorimotor synchronization (Nozaradan et al., 2016), recording brain activity by the means of the EEG during a passive listening task and subsequently performing the behavioral task. The authors detected entrained cortical activity on the frontocentral cluster of electrodes where auditory responses are typically detected, hypothesizing that SSEPs amplitudes would predict behavioral measures of overt entrainment. Interestingly enough, a dissociation emerged in the correlations between their measure of neural entrainment and behavioral accuracy, when compared to behavioral consistency. Specifically, the amplitude of SSEPs was related to mean asynchrony (accuracy) rather than to the resultant vector length (consistency), suggesting that the two are supported by distinct neural mechanisms when processing the beat of an auditory rhythm. In the scenario we proposed, with the goal of relating neural entrainment to 'online' dynamics of behavior, we identified a lateralized component plausibly related to primary sensorimotor areas, whose stability index

happens to be related to both behavioral accuracy and consistency. Our finding is arguably not in contradiction with previous evidence, but rather complementary.

One may argue that the correlation we found between the stability index and the resultant vector length could be spurious, a sort of epiphenomenon entirely explainable by afferent proprioceptive feedback. Following this argument, stable rhythmic behavior could produce steady responses in primary somatosensory areas (Bourguignon, Piitulainen, De Tiège, Jousmäki, & Hari, 2015; Piitulainen, Bourguignon, De Tiège, Hari, & Jousmäki, 2013). Our task cannot exclude the possibility that such afferent components lead to a spurious correlation between the stability index and resultant vector length, which quantifies behavioral consistency. Nevertheless, such interpretation cannot explain our crucial finding that the stability index also correlated with behavioral accuracy. A more stable entrained component was associated with smaller synchronization errors, as quantified by measures of asynchrony and relative phase. We argue that our results rather align with evidence that motor cortices play a critical role in supporting auditory perception and prediction (Assaneo, Rimmele, Sanz Perl, & Poeppel, 2021; Takako Fujioka et al., 2012; Morillon & Baillet, 2017), since we showed that the more stable the entrained component is, the smaller the error relative to the beat. Within the limits of our auditory-motor task, we pick up on the notion of active sampling (Morillon et al., 2019), to propose that entrainment dynamics driven by the motor system seems to play an active role in the predictive mechanism of error minimization underpinning auditory-motor coupling (Vuust & Witek, 2014). However, with the current experimental design, we cannot rule out that distinct motor, sensory and cognitive processes were to some extent mixed in the entrained component. This represents an important limitation of the present study. A finer disentanglement of the neural processes underlying entrainment should be addressed by future work, with dedicated experimental designs.

With our work, we thus contributed methodologically to the investigation on neural entrainment. Our method consists of extracting the oscillatory component in the EEG signal which is maximally entrained to a rhythmic auditory stimulus, and subsequently quantifying the stability of fluctuations over time. The impact of this contribution has valuable prospects within the domain of neurological rehabilitation. In previous work, we have investigated motor and auditory entrainment in participants with multiple sclerosis and healthy controls. Specifically, previous studies have analyzed behavioral time-series by means of detrended fluctuation analysis (DFA) (Moumdjian et al., 2020). Differences in gait dynamics were attributed to the process of error-correction minimization, which are required to dynamically interact with continuous and discrete auditory structures typically present in music and metronomes, respectively. Although clinically relevant, complementing such studies with neural outcome measures such as the stability index

would allow to explain the process of error-correction minimization further, at the level of neural dynamics. Such prospect has a strong indication to optimize the individualized rehabilitation procedure.

In conclusion, our approach can be used for better understanding the dynamics of an entrained system over time. While the stability index provides a global neural outcome measure correlated with the overall synchronization performance, the instantaneous frequency time-series can offer a more fine-grained picture on the dynamics of neural entrainment. Neural and behavioral measurements can be complemented within a comparative setting between heathy population and neuropathological models, offering the possibility to dissociate neural mechanisms based on a mapping of selective lesions. Such neuropathological models can be recruited through studies conducted on participants with neurological diseases, where components of cognitive, motor or perceptual functions can be isolated. For instance, cerebellar lesions cover particular interest given the role of cerebellum in encoding the timing of events at the micro-timing scale (R. B. Ivry & Keele, 1989; R. B. Ivry, Keele, & Diener, 1988; Richard B. Ivry & Schlerf, 2008), and given that their neural entrainment to auditory rhythms is selectively compromised at faster tempi (Nozaradan, Schwartze, Obermeier, & Kotz, 2017). Respectively, this unfolding of observations would expand the knowledge of the complex dynamic interaction when entraining motor and auditory systems to one another. In turn, it would pave ways towards the development of state-of-the-art approaches within the domain of neurological rehabilitation.

Author contribution

M.R. was responsible for study conceptualization, data collection, processing and analysis, and writing of the manuscript. M.L. was responsible for data interpretation and manuscript writing. L.M. was responsible for study conceptualization, data interpretation and writing of the manuscript.

Funding

The present study was funded by Bijzonder Onderzoeksfonds (BOF) from Ghent University (Belgium).

Conflict of interest

Authors declare no conflict of interest.

Acknowledgements

We would like to acknowledge Mr. Ivan Schepers, for his technical support towards developing the tapping pad hardware.

7

Neural entrainment underpins sensorimotor synchronization to dynamic rhythmic stimuli

Mattia Rosso [A,B], Bart Moens [A], Marc Leman [A], Lousin Moumdjian [A,C,D]

[A.] IPEM - Institute for Systematic Musicology; Ghent University; Ghent, Flanders, 9000; Belgium.

[B.] PSITEC - Psychologie: Interactions, Temps, Emotions, Cognition - ULR 4072; University of Lille; Lille, Hauts-de-France, 59650; France.

[C.] UMSC Hasselt-Pelt, Belgium

[D.] REVAL Rehabilitation Research Center, Faculty of Rehabilitation Sciences, Hasselt University

Rosso, M., Moens, B., Leman, M., & Moumdjian, L. (2023). Neural entrainment underpins sensorimotor synchronization to dynamic rhythmic stimuli. NeuroImage, 120226.

Introduction

Humans exhibit a natural inclination to synchronize their movement with rhythmic signals in the environment. The phenomenon can be observed across a range of contexts such as music playing and dance (Clayton et al., 2020; Marc Leman, 2016), sports (E. E. A. Cohen, Ejsmond-Frey, Knight, & Dunbar, 2010), joint rhythmic tasks (Rosso et al., 2022; Rosso, Maes, et al., 2021) and verbal communication (Richardson et al., 2008; Shockley et al., 2003). How does the human brain represent an external rhythm, how does it track it over time, and how does it temporally match motor behavior to it? These questions cover nowadays exceptional interest in neuroscience. Evidence converges towards sensorimotor entrainment, namely the alignment of motor and sensory rhythms at the neural level, as a plausible answer.

Neural entrainment is defined as unidirectional synchronization of neural oscillations to an external rhythmic stimulus (Haegens & Zion Golumbic, 2018; Lakatos et al., 2019), and comes with the assumption of endogenous oscillatory activity in the brain which can be driven towards a state of phase- and frequency-locking. From the viewpoint of perception, oscillations reflect changes in the weight of sensory inputs via rhythmic fluctuations in neuronal excitability (Buzsáki & Draguhn, 2004; Fries, 2005; Lakatos et al., 2005). Their entrainment to rhythmic events is thought to subserve the selection of relevant information and to reduce the interference of competing input streams (Lakatos et al., 2019). Perceptual representations are enhanced via low-frequency rhythmic fluctuations of sensory gain (Obleser & Kayser, 2019), resulting in what is considered to be a 'rhythmic mode' of attention (Obleser, Henry, & Lakatos, 2017; Rosso et al., 2022; Schroeder & Lakatos, 2009; Schroeder et al., 2010; Zoefel & VanRullen, 2015). Noteworthy, low-frequency oscillations originating from cortical motor areas (Morillon & Baillet, 2017; Morillon, Hackett, Kajikawa, & Schroeder, 2015; Morillon, Schroeder, & Wyart, 2014) are thought to guide selection of environmental information and active sampling implemented in motoric routines (Schroeder & Lakatos, 2009), which leads to perceptual enhancement of attended stimuli (Chemin, Mouraux, & Nozaradan, 2014; Morillon et al., 2014; Park, Ince, Schyns, Thut, & Gross, 2015; Rimmele et al., 2018; Schroeder et al., 2010). Thus, the engagement of the motor system enhances and stabilizes the internal representation of environmental rhythms (Kliger Amrani & Zion Golumbic, 2022) by scaffolding the prediction of incoming sensory events (Arnal & Giraud, 2012; Morillon et al., 2019, 2015, 2014; Morillon & Baillet, 2017; Rimmele et al., 2018; Rosso et al., 2022). A network of cortical and subcortical motor areas was recently put forward as responsible for beat-based time-keeping, by dynamically tracking the phase of the stimulus cycle and enabling overt motor alignment (Cannon & Patel, 2021).

The rationale of current state-of-the-art approaches to quantify neural entrainment is that the frequencies of a rhythmic stimulation can be tagged in the power spectrum of electrophysiological timeseries (Nozaradan et al., 2011). In practice, 'frequency tagging' transforms the brain signal into the frequency domain, quantifying the power of stimulus-related frequencies over the whole spectrum. The observation that target frequencies dominate the spectrum has been commonly taken as evidence for the underlying entrainment of neural oscillations (e.g., (Lenc et al., 2018; Nozaradan et al., 2011, 2012)), which is not exempt from critiques. The most evident of these highlights that periodic stimulation elicits a train of time-locked transient responses in the brain, resulting in prominent peaks in the power spectrum at the related frequencies (Novembre & Iannetti, 2018). Furthermore, allocating attentional resources to a predictable stimulus results in top-down modulation of the evoked response amplitude (Breska & Deouell, 2017; Legrain, Iannetti, Plaghki, & Mouraux, 2011; Nobre & van Ede, 2018) and, in turn, of power at the related frequency. Periodic evoked responses hinder the measurement of real neural entrainment, represent a critical confound for the empirical investigation with non-invasive electrophysiology, and pose a major methodological challenge for the researchers in the field (Haegens & Zion Golumbic, 2018).

Fundamentally, the entrainment of two oscillatory signals cannot be described by their spectral profile alone, because this is not a test for the underlying oscillatory process (Obleser & Kayser, 2019). Given that the temporal structure of a signal strictly depends on the phase of its oscillatory components (Rajendran & Schnupp, 2019), it is not possible to make inferences on local oscillatory dynamics of the signal if phase information is neglected. Neural entrainment is just one possible process leading to the match of two spectral profiles, and therefore frequency tagging is not a sensitive method to its realization. Entrainment in the strict sense is based on the phase of the signal, and is in principle dissociated from its amplitude. Therefore, the process should be defined in terms of changes in frequency over time (Rosenblum, Pikovsky, Kurths, Schäfer, & Tass, 2001) rather than being inferred by the power spectrum. It is the unidirectional process that leads to a state of phase-locking with a driving force (Lakatos et al., 2019), not the phase-locked state itself (Rosenblum et al., 2001).

Motivated by these limitations, we previously proposed the *Stability Index (SI)* as a measure to quantify neural entrainment from electroenchephalograpy (EEG) recordings of healthy participants engaged in finger-tapping to a steady auditory metronome (Rosso, Leman, et al., 2021). In our work, we extracted from the EEG signal a component attuned to the stimulation frequency, and computed the *SI* based on the fluctuations of the component's frequency over time. Critically, we reported significant correlations between the *SI* and behavioral measures of synchronization: the more stable the neural

component, the more stable and more accurate the synchronization performance. We argued that, in contrast to previous amplitude-based approaches, the measure explicitly captures the dynamic phase-adjustment of entrained neural oscillations. The putative entrained component would adaptively speed up and slow down, fluctuating around the target center frequency to reach stable synchronization over time (Rosso, Leman, et al., 2021). However, one main limitation of our previous work is that we could only provide a global measure of these fluctuations for a given time window, and correlate it to the global behavioral performance. Further inferences on the dynamics of the frequency adjustment over time were not possible, due to the fact that we did not have an experimental paradigm to induce them in a controlled fashion.

With these limitations in mind, we hereby present the '*event-related frequency adjustment*' (ERFA) as a novel experimental paradigm for investigating neural and behavioral entrainment with auditory rhythmic stimuli. A way to induce controlled frequency adjustment is to manipulate the frequency of an auditory metronome while experimental subjects are attempting to synchronize their finger-taps to it. The event discloses a post-perturbation window wherein the entrained component is expected to adjust its frequency according to the stimulus dynamics, and correct the error to return to a stable state. In finger-tapping studies, error-correction is traditionally investigated with tempo-changes and phase-shifts, in both positive and negative directions (for a review, see (Bruno H. Repp & Su, 2013). We implemented the former as a step change of +/- 10% from the baseline frequency (1.67 Hz), and the letter as a +/- 90° shift of the metronome beat along its cycle. Given that control mechanisms are thought to underpin error correction depending on the nature and the direction of perturbations (L. Bavassi, Kamienkowski, Sigman, & Laje, 2017; Jantzen, Ratcliff, & Jantzen, 2018; Praamstra, Turgeon, Hesse, Wing, & Perryer, 2003; B. H. Repp, 2001b, 2001a), these variables were expected to reveal different underlying oscillatory dynamics.

The ERFA curves, which constitute the neural measure within our experimental paradigm, are computed in three main blocks: 1) *attunement*: from the continuous multivariate EEG signal, a single component maximally attuned to the metronome's frequency is extracted via spatial filtering (Michael X. Cohen & Gulbinaite, 2017; Rosso, Leman, et al., 2021); 2) *instantaneous frequency*: changes in frequency over time are computed based on the rate of change of the component's phase (Michael X. Cohen, 2014; Rosenblum et al., 2001); 3) *event-based segmentation*: instantaneous frequency responses are time-locked to the perturbations, and aggregated by perturbation type and direction. The *attunement* block is carried out separately on two perturbation-free periods, while actively listening to the stimuli or tapping along, which allows to untangle a perceptual and a sensorimotor

entrained component within the same rhythmic task, and assess their relative contributions in adjusting to perturbations.

Our aim was to use ERFA to track neural entrainment dynamics within the post-perturbation windows, and to model them as a function of time. We hypothesized that the instantaneous frequency response would track the stimulus dynamics across perturbation types and directions, and that the active engagement in the behavioral task would boost the entrainment of the sensorimotor component in the brain signal as compared to the perceptual component. *Figure 7.1* provides a graphical representation of paradigm, trial structure and expected results. Procedures are explained in detail in the *Materials and methods* section.

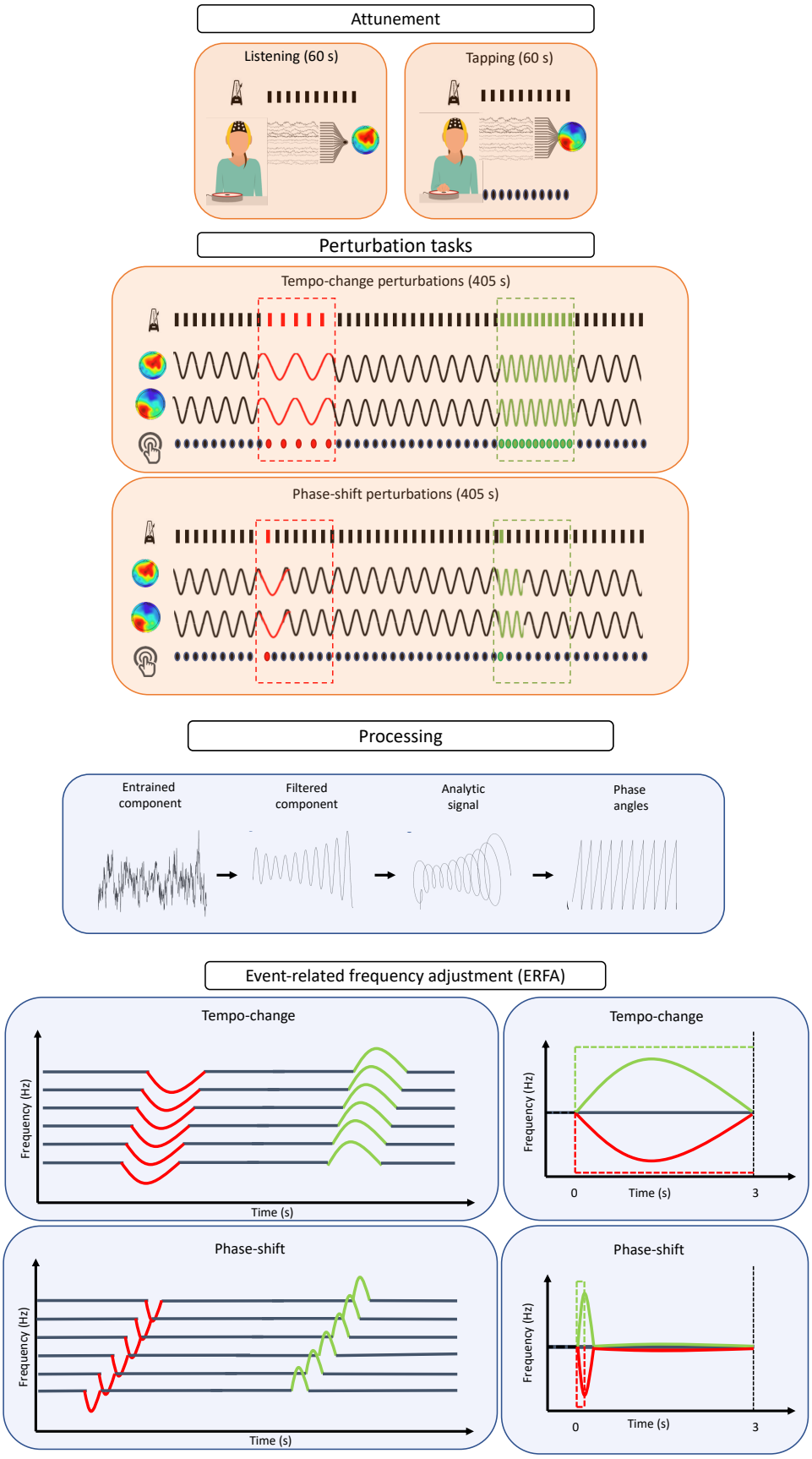

Attunement
Listening (60 s)
Tapping (60 s)
Perturbation tasks
Tempo-change perturbations (405 s)
Phase-shift perturbations (405 s)
Processing
Entrained component
Filtered component
Analytic signal
Phase angles
Event-related frequency adjustment (ERFA)
Tempo-change
Frequency (Hz)
Time (s)
Tempo-change
Frequency (Hz)
0
Time (s)
3
Phase-shift
Frequency (Hz)
Time (s)
Phase-shift
Frequency (Hz)
0
Time (s)
3

Figure 7. 1. ERFA paradigm workflow. Orange boxes provide a context of experimental tasks, blue boxes represent data processing. The approach started with an attunement phase, where participants were exposed to a perturbation-free auditory metronome during one minute. The purpose was to induce a stable oscillation attuned to the stimulation frequency, facilitating its separation from the broadband multivariate EEG signal via generalized eigendecomposition (GED) (Michael X. Cohen, 2022; Michael X. Cohen & Gulbinaite, 2017). Participants underwent this period of steady stimulation under distinct conditions of active listening and finger-tapping: using these two datasets, we designed for every participant two different spatial filters to extract a purely perceptual (without movement) and a sensorimotor (with movement) entrained components, respectively. After 60 seconds of finger-tapping, metronomes started to be perturbated by introducing unpredictable tempo-changes or phase-shifts during 405 seconds, depending on the experimental condition. During these perturbation tasks, participants were instructed to synchronize their finger-taps to the metronome, therefore following the stimulus dynamics. The spatial filters designed based on the attunement phase were applied to the EEG signal recorded during the tasks, allowing us to track the dynamic changes in the frequency of both perceptual and sensorimotor oscillatory components. Examples of negative (red) and positive (green) perturbation windows are represented for tempo-changes and phase-shifts, illustrating how the metronomes' onset were manipulated and how neural components and finger-taps were expected to entrain. These figures are presented as conceptual examples for illustrative purposes only and are not meant to be realistic. In order to provide a measure for the neural and behavioral adjustments, the following processing pipeline based on Rosso et al. (Rosso, Leman, et al., 2021) was applied. Entrained components were narrow-band filtered around the metronome's center frequency and Hilbert-transformed to produce analytic signals and extract phase timeseries. These were then unwrapped to prevent phase resets, differenced and scaled to Hz to produce instantaneous frequency timeseries, namely an estimate of the oscillation's frequency at every timepoint. The processing up to this stage was performed on the continuous EEG recording. Instantaneous frequency timeseries were eventually segmented based on perturbation windows, and aggregated by perturbation type and direction. The average of the event-based segments provided the event-related frequency adjustment (ERFA) curves as our outcome of interest, which were expected to follow the stimulus dynamics as expressed in terms of instantaneous frequency. In the rightmost boxes of the figure, we show how we hypothesized the ERFA curves would look like. We expected a gradual and sustained response to tempo-changes (top box) and a transient response to phase-shifts (bottom box), tracking the stimulus dynamic and direction (represented as dashed lines).

Materials and Methods

1. Participants

Twenty (N = 20) right-handed healthy participants took part in the experiment (11 females, 9 males; mean age = 32.8 years, std = 6.2 years). All participants had normal hearing and normal or corrected-to-normal vision; none reported any history of major medical, psychiatric or neurological conditions; none reported to be a professional musician. Upon arrival in the laboratory, they received extensive briefing of the experimental procedure and signed the informed consent form. Following the administration of the Edinburgh Handedness Inventory (Oldfield, 1971) and a questionnaire to collect information on demographics and musical training, the experimenter proceeded to prepare the EEG equipment.

The study was approved by the Ethics Committee of Ghent University (Faculty of Arts and Philosophy) and informed written consent was obtained from each participant. A 20€ coupon was given to all participants as economic compensation for their time.

2. Experimental tasks

The experiment consisted of one listening task without perturbations and two finger-tapping tasks with perturbations, performed in a fully-randomized order. Two different perturbation types were used as separate experimental conditions, these were tempo-changes and phase-shifts (see *Figure 7.1*). The listening task lasted 60 seconds, while both finger-tapping tasks lasted 465 seconds, of which the first 60 seconds were free of perturbations and the following 405 seconds contained perturbations.

The listening task was used to collect EEG data in absence of movement and design a spatial filter, to extract the perceptual component maximally attuned to the auditory stimulus. Stimulation consisted of a metronome set at 100 Bpm (1.67 Hz), which was chosen as optimal rate for sensorimotor synchronization (Kliger Amrani & Zion Golumbic, 2022; London, 2012; Zalta, Petkoski, & Morillon, 2020). To verify that participants were actively listening, they were instructed that the regular rhythm could be disrupted by a phase-shift at any moment, and they would have to tap their finger on the pad as fast as possible upon detection. One single perturbation was presented after 60 seconds of steady rhythm, at a random moment within a 5 second time window, and was successfully detected by all participants but one. The importance of not moving any body part during

the listening was stressed, as well as staring at a black dot on the center of the pad to minimize eye-movement artifacts in the EEG recordings.

In finger-tapping tasks, participants were instructed to synchronize their finger-taps to the metronome all the time, and to keep trying to synchronize even in case they would find themselves out of sync. The importance of minimizing body and head movements during the task was stressed, as well as staring at a black spot (i.e. a fixation point) painted on the tapping pad. Both measures were taken to minimize contamination of EEG recordings. Immediately before the start of the experiment, participants were given a simple demonstration by the experimenter and familiarized themselves with the finger-tapping to the non-perturbated metronome until they felt confident with the procedure.

Each finger-tapping task lasted a total of 465 seconds, the first 60 of which were free from perturbations. Such steady period was used to design a spatial filter to extract the sensorimotor component maximally attuned to the metronome, based on EEG data collected in condition of overt movement. From that moment onwards, perturbations were presented at random times every 5 to 15 seconds: perturbation onsets were therefore unpredictable while participants had enough time to re-stabilize behavioral and neural responses. A total of 20 negative perturbations, 20 positive perturbations, and 40 perturbation-free windows were available to analyze per participant within a time frame of 405 seconds. The order of perturbations' direction was randomized within finger-tapping condition, whilst tempo-change and phase-shift were presented in separate experimental conditions. Tempo-changes consisted of a +/- 10% step change with respect to the baseline tempo (100 bpm, 1.667 Hz), sustained for 3 seconds before a second step change back to the baseline. Phase-shifts consisted of a discrete +/- 90° shift on the beat cycle, perceived as a shorter or a longer inter-beat interval, respectively (see *Figure 7.1*). The order of the individual conditions (60 seconds of perturbation-free listening, 465 seconds of finger-tapping with tempo-changes, 465 seconds of finger-tapping with phase-shifts) was fully randomized across participants.

Besides the experimental conditions hereby presented, participants also performed the same tasks with different kinds of ecological musical stimuli. However, such extra conditions were meant to answer research questions not relevant to the present paper, hence will be addressed in future work.

3. Experimental apparatus

Participants were sitting on a comfortable armrest chair, and a circular pad was placed on the table in front of them for recording finger-tapping responses. Stimuli were presented

via DefenderShield® air-tube earplugs. Volume was adjusted before the beginning of the experiment, to make sure that stimuli were clearly audible without creating any discomfort. A circular tapping pad, containing a strain gauge pressure sensor, was used to detect finger-tapping onsets with a 1 ms resolution. A folded towel was placed underneath the pad, so that no auditory feedback from the finger-taps was perceivable. The pressure sensor was connected to a Teensy 3.2 microcontroller, which worked as serial/MIDI hub for data logging and communication between the stimulation computer and the EEG recording computer. Upon occurrence of events of interest (i.e., perturbations), a TTL trigger was sent from the Teensy microcontroller to the EEG amplifier via BNC connection, granting the alignment between behavioral and neural timeseries. The EEG signals were recorded with an ANT-Neuro eego™mylab system at a sampling rate of 1 kHz. Each participant was equipped with an EEG headset (64-channel waveguard™original with Ag/AgCl electrodes). Recordings were performed with a referential montage, with 'CPz' being the reference for all electrodes.

The stimuli sequence was played back by specifically designed software developed in Max MSP 8 (Cycling '74, USA), running on the stimulation computer (Windows 10, Intel core i7 8th gen, Focusrite Rednet PCIExpress ASIO low-latency soundcard). Prior to each trial, a randomized balanced list of perturbations was generated and inspected by the experimenter. A pre-generated .wav file, containing a non-perturbed sequence of metronome ticks at 1.667 Hz (600ms inter-beat intervals), was played back and manipulated in real-time based on the perturbations list and timing configuration (i.e., initial perturbation-free period and minimum and maximum time between perturbations). MIDI events were generated when a beat was perceived by the listener and when a perturbation occurred. These MIDI events were logged using the Teensy 3.2 microcontroller, along with the finger-tapping timestamps.

4. Data analysis

The data processing pipelines were implemented in Matlab 2019a (MathWorks Inc., USA). Statistical modelling was carried out in R Studio version 4.0.3 (R Core Team), using the lme4 package (Bates et al., 2014) for model fitting. Participant #6 was excluded from the analysis, due to a technical issue during data acquisition resulted in the loss of EEG data.

4.1. Behavioral data processing

Finger-taps were processed in order to return timeseries aligned to the neural ERFA, expressed in the same unit of measure (i.e., Hz over ms). Whenever a timestamp was followed by a second timestamp with < 350ms interval, it was considered as a false positive and therefore removed (false positives could occasionally be recorded when a participant pushed the tapping pad for too long or accidentally laid the hand on it). The intervals between the remaining timestamps were then linearly interpolated from 0 to 1 at 1 kHz sampling rate; the resulting ramp wave was scaled to 2π, providing an estimate of the finger-taps phase with a temporal resolution of 1 ms. Instantaneous frequency timeseries were computed as the first derivative of the unwrapped phase angles time series (Boashash, 1992), and converted in Hz as indicated in (Michael X. Cohen, 2014):

$$Hz_t = \frac{s(\phi_t - \phi_{t-1})}{2\pi}$$

where *s* indicates the data sampling rate and ϕ_t indicates the (unwrapped) phase angle at time *t*. Unwrapping was necessary in order to remove discontinuities in the timeseries caused by phase resets.

4.2. Neural data processing

4.2.1. Pre-processing

The pre-processing pipeline was realized integrating functions from the *Fieldtrip* toolbox (Oostenveld et al., 2011) for Matlab. Bad channels were identified by visually inspecting the raw timeseries in combination with the distribution of variance across channels. The recording was re-referenced to the average of all electrodes after rejection, to avoid noise leakage into the common average. 3.23 bad channels per participant were removed on average (std = 1.95). A sixth-order Butterworth high-pass filter with 1 Hz cut-off was applied to the raw recordings to remove slow drifts. We show in *Supplementary material 7. A* that, for these parameters, the high-pass filter does not influence the oscillatory dynamics of interest in the present work. A low-pass sixth-order Butterworth filter with 40 Hz cut-off was applied to remove high-frequency muscular activity. A fourth-order notch filter centered at 50 Hz was applied to remove power-line noise up to the 3rd harmonic.

Subsequently, independent component analysis (ICA) was conducted and used to remove stereotyped artifacts by means of visual inspection of topographical maps and timeseries of component activation, as implemented in the 'runica' *Fieldtrip* algorithm. The reference 'CPz' and the bad channels' timeseries were excluded from the input matrix. Under optimal conditions, removal was limited to those components which exhibited the stereotypical frontal distribution generated by blinks and lateral eye movements, or bilateral temporo-mastoidal distribution with periodic peaks in the activation timeseries plausibly generated by heart beats. Extra components were removed in instances where recurrent artifacts with clearly abnormal amplitude were detected. 5.03 artifactual components per participant were identified and removed on average (std = 4.36). The dataset was inspected prior to ICA decomposition and following ICA back-projection to assess the quality of the artifact removal. Special attention was given to the frontal clusters of electrodes maximally contaminated by eye-related artifacts. Rejected bad channels were reconstructed after artifact removal, by computing the average of activity from neighboring electrodes indicated by the template provided by ANT-Neuro for 64-channel waveguard™original caps. No segmentation in epochs was performed up to this point, since continuous recording was needed for performing source separation.

4.2.2. Source separation

Generalized eigendecomposition (GED) (Michael X. Cohen, 2022) was used to extract a perceptual and the sensorimotor components attuned to the stimulation frequency. The procedure described below is the same for both components, with the only difference that the respective inputs were 60 seconds of data recorded during the listening task, and 60 seconds of data recorded during the 60 seconds of finger-tapping in absence of perturbations. This allowed to design two separate spatial filters, which were subsequently applied to the data recorded during the 465 seconds of finger-tapping in presence of perturbations (see *Figure 7.1*).

As first described in the context of source separation for rhythmic entrainment (Michael X. Cohen & Gulbinaite, 2017), GED allows to avoid channel selection bias while optimizing the signal-to-noise ratio between the entrained component and the broadband neural activity. The technique consists of a spatial filter to separate sources and reduce data dimensionality, guided by some criteria. In this case, the criterion is the attunement to the stimulation frequency. Dimensionality was reduced by computing the weighted average of the timeseries from all 64 channels, where the set of vectors W (weights) was calculated by solving the following equation:

$$RW\Lambda = SW$$

where S is the covariance matrix of the narrow-band signal; R is the reference covariance matrix of the broad-band signal; Λ is the set of eigenvalues. GED identifies eigenvectors W that best separate the signal ('S') covariance from the reference ('R') covariance matrix. The eigenvector associated to the largest eigenvalue was taken as a spatial filter, transposed and multiplied by the broadband data matrix to reconstruct the time series of our target entrained component.

Given we were explicitly looking for frequency fluctuations, our narrow-band filter needed to be wide enough to leave room for fluctuations around the target frequency. We designed our filter as a gaussian function in the frequency domain, with center at 1.667 Hz and a width of 0.3 Hz at half of the maximum gain (Rosso, Leman, et al., 2021). We then filtered the broadband data at all channels via spectral multiplication between broadband signal and wavelet kernel in the frequency domain, and transformed the resulting narrow-band signal back in the time domain with inverse fast Fourier transform.

The S covariance matrix was computed from the narrowband signal, the R covariance matrix was here computed from the broadband multivariate signal. In this regard, the choice aligns with the approach presented in (Rosso, Leman, et al., 2021) rather than (Michael X. Cohen & Gulbinaite, 2017), because the former is optimized for low target frequencies < 2 Hz. Covariance matrices were computed within 600 ms time windows locked to the finger-taps onsets, and grand-average S and R covariance matrices were computed. Matrices whose z-normalized Euclidean distance from the grand-average exceeded the 2.23 z-scores (corresponding to a probability of 0.01) where removed, and the grand-average S and R recalculated free of transient artifactual activity. 1% regularization was applied to R.

In previous work (Rosso et al., 2022; Rosso, Leman, et al., 2021), we implemented an optimization of GED based on a macro-selection of regions of interest, justified by experimental design and some prior assumptions. In this study, however, we made no prior assumptions on scalp distribution and therefore inputted timeseries from all channels, allowing the data collected to drive the source separation in the different conditions ('listening' and 'finger-tapping').

4.2.3. Instantaneous frequency

The same Gaussian filter used for GED was applied to the perceptual and to the sensorimotor components (center at 1.667 Hz and 0.3 Hz width at half maximum) to extract reliable phase timeseries from the analytic signal (Rosenblum et al., 2001). These were extracted by means of Hilbert transform. In order to remove discontinuities caused by phase resets, the timeseries were unwrapped, differenced and finally converted to Hz (Michael X. Cohen, 2014):

$$Hz_t = \frac{s(\phi_t - \phi_{t-1})}{2\pi}$$

The resulting instantaneous frequency timeseries were smoothed with a sliding moving median (window width of 400 samples), to remove transient artifactual activity that may distort the phase timeseries (Michael X. Cohen, 2014).

It should be to pointed out that, in electrophysiological data, the estimation of instantaneous frequency of oscillatory activity is sensitive to the aperiodic 1/f component of the spectrum (Donoghue et al., 2020). This feature can result in bias towards lower frequencies under certain conditions. Samaha and Cohen (Samaha & Cohen, 2022) demonstrated that a low periodic/aperiodic spectral power ratio, in combination with broad filter width, is problematic for reliable estimations. We highlight that our spatial filter application was specifically aimed at maximizing the target narrowband energy relative to the broadband energy (Michael X. Cohen & Gulbinaite, 2017; Haufe, Dähne, & Nikulin, 2014; Rosso, Leman, et al., 2021), resulting in outstanding spectral peaks indicating prominent oscillatory activity (see *Figure 7.2*). This, along with our conservative choice of a narrow filter, complies with the good practices to reducing the slope of the 1/f, and therefore increases the reliability of the estimated instantaneous frequency (Samaha & Cohen, 2022).

4.2.4. Event-related frequency adjustment (ERFA)

In order to analyze neural and behavioral responses to perturbations within the same framework, the following procedure was applied to all instantaneous frequency timeseries to compute the related instances of ERFAs: one for the finger-tapping (behavioral), one for the EEG perceptual component (neural) and one for the EEG sensorimotor component

(neural). The approach was inspired by event-related potentials (ERPs) (Luck, 2014) and shares with it most of the features presented in this paragraph. The crucial difference is that ERFAs express frequency (y-axis) over time (x-axis), whilst ERPs express amplitude (y-axis) over time (x-axis). Furthermore, the pre-stimulus baseline of ERFAs requires steady rhythmic stimulation to provide a stable baseline frequency level.

Perturbation onsets were identified in the instantaneous frequency timeseries based on the timestamp logs. Time windows were defined as the time span from -500 to +3000 ms with respect to the perturbation onset, aggregated per perturbation type and direction, baseline-normalized and averaged. Baseline normalization was performed by subtracting the average of the 500 to 0 ms interval from the rest of the ERFA, dividing by the target stimulation frequency (i.e., 1.667 Hz) and multiplied by 100. The resulting ERFA is expressed in percent change respect to the baseline stimulation frequency. For every participant, 19 ERFA curves were aggregated per perturbation type and direction, and the average response was computed. The responses to negative perturbations were sign-flipped in order to avoid trivial results at statistical comparisons with positive perturbations (Jantzen et al., 2018). Finally, the ERFA computation was repeated shifting all time-windows by -400 ms along the instantaneous frequency timeseries, namely into perturbation-free periods. This provided for every participant a baseline for further statistical comparisons.

5. Statistical modelling

ERFAs to tempo-change perturbations were modelled via polynomial fitting to the curves, down-sampled by a factor of 10 for computational feasibility. For phase-shifts, the discrete integral of the ERFA from 0 to 1500 ms interval was computed via trapezoidal method with unit spacing, using the *trapz()* Matlab function. This returned signed areas under the curves, which we interpreted as a measure of entity of the error correction sensitive to the direction of the corrective response. Baseline normalization is crucial in order to obtain signed areas, as data need to be zero-centered.

For behavioral ERFAs only, reaction times were computed as the x-coordinate of the inflexion point of sigmoid function which best fitted the ERFA from 0 to 800 ms. The fitting was performed with the *sigm_fit()* Matlab function, defining no fixed parameters. For the neural ERFAs, the effects on the response variables of interest were tested in 3x2 factorial designs with Direction (Negative, Null, Positive) and Component (Perceptual, Sensorimotor) as factors. For the behavioral ERFAs, the effects on the response variables of interest were tested in 3x1 factorial design with Direction (Negative, Null, Positive) as

the only factor. When pairwise comparisons between Negative and Positive directions were performed after omnibus tests, Bonferonni correction was applied for a significance level of $\alpha = 0.05$.

5.1. Orthogonal polynomials (for tempo-change)

For tempo-changes, growth curve analysis (Mirman, 2017; Rosso, Maes, et al., 2021) was used to model the change in instantaneous frequency of the neural entrained components (i.e., the ERFAs) within 3-seconds perturbated windows (i.e., metronomes at +/- 10% of the baseline frequency). Parameters of categorical predictors were estimated relative to Null Direction and Perceptual component, respectively. The rationale of such factor leveling is that no systematic change in instantaneous frequency was expected in Null trials, resulting in the flattening of random fluctuations when averaging over trials. As for the Component, we intended to test the significance of additive motor processing in the Sensorimotor component during entrainment, with respect to purely Perceptual processing.

The order of the polynomial is ideally chosen based on hypothesis and on the nature of the data, should be confirmed by the empirical data, and should allow a straightforward interpretation of the effects (Mirman, 2017). A full model based on orthogonal polynomials includes all terms up to the chosen order, which in this case was 2: a linear and a quadratic term were sufficient to capture the effect of our manipulation. In a framework of polynomial fitting, a flat line can be considered a pure intercept and a 'zero-order' polynomial, in the sense that it exhibits zero-change in any direction. When instantaneous frequency is time-invariant, it indicates a steady oscillation within a given time window (e.g., the behavior of a non-perturbated metronome). The intercept of our model provides hence valuable information on the average frequency within the perturbated window, and how it changes across the factors' levels: significant main effects of categorical predictors indicate global differences across experimental conditions, independently from the temporal profile of the response variable. On the other hand, significant interaction effects on the polynomial terms indicate that instantaneous frequency in systematically modulated by the temporal dynamics of the perturbation. The parameter estimate of the linear component (1st order) corresponds to the slope of the line and the consequent shift in the vertex for the parabola, while the quadratic component (2nd order) corresponds to the parabolic curvature. In sum, intercepts indicate average frequency; polynomial terms of Time are used to model instantaneous changes in frequency. Our model also included random effects of Subject on all polynomial

terms, and their interactions with the factors: the random effects structure was maximized in order to minimize false alarm rates without substantial loss of power (Barr et al., 2013).

Here, the formulas for the models fitted to neural and behavioral ERFAs, respectively (in R syntax):

$$nERFA \sim (Time + Time^2) * Component * Direction + (Time + Time^2 \mid Subject) + (Time + Time^2 \mid Subject{:}Component{:}Direction)$$

$$bERFA \sim (Time + Time^2) * Direction + (Time + Time^2 \mid Subject) + (Time + Time^2 \mid Subject{:}Direction)$$

5.2. Integrals (for phase-shifts)

For phase shifts, we deemed a model based on integrals to bemore suited and parsimonious as compared to polynomials. This was motivated by the observation that a parabolic curve did not provide a good fit for transient responses elicited by phase shifts. We therefore opted for a method insensitive to the particular curve-shape, quantifying instead the area under the curve within a shorter time window (1500 ms) following the perturbation onset. The adoption of integrals in the context of event-related neural responses was validated in the ERP literature (Luck, 2014), and in principle equally valid in the context of these ERFAs given the comparable post-stimulus window sizes and a response more localized in time, in contrast to the more sustained responses to tempo-changes.

The integrals provided us with a measure of ERFA magnitude, which is sensitive to the direction of the response since the areas are signed with respect to the 0% change baseline. Noteworthy, we verified on the data logged from the metronomes that the expected absolute area underneath positive and negative phase-shifts were equal irrespective of the difference in shape. Here are the formulas for the models fitted to neural and behavioral ERFAs, respectively, plus the behavioral reaction times (in R syntax):

$$nERFA\_integral \sim Component * Direction + (1 \mid Subject)$$

$$bERFA\_integral \sim Direction + (1 \mid Subject)$$

$$bERFA\_rt \sim Direction + (1 \mid Subjct)$$

Data and scripts are available upon request to the authors with a formal data sharing agreement, in line with the conditions of the local ethics committee which approved the present study.

Results

Event-related frequency adjustments (ERFAs), computed as the instantaneous frequency response within the post-perturbation window, are the object of our analyses. Neural and behavioral ERFAs were computed from EEG and finger-tapping timeseries, respectively. Two distinct approaches were used to model the ERFAs in response to the tempo-change and phase-shift perturbations. For both neural and behavioral responses, orthogonal polynomials (Mirman, 2017) were used to model tempo-change, and integrals (Luck, 2014) were used to quantify the entity of phase-shifts. ERFAs in response to all negative perturbations were sign-flipped, in order to make the responses directly comparable with the positive counterparts (Jantzen et al., 2018). Whereas in the literature stimulus rates are often expressed in milliseconds to describe the duration of inter-stimulus intervals, ERFAs will be consistently expressed in Hz, or % change with respect to a baseline frequency. Therefore, increasing values indicate that the oscillatory component is speeding up, and decreasing values indicate that the oscillatory component is slowing down.

Below, an assessment of the EEG source separation is provided, followed by the results of our models presented per perturbation type.

1. EEG source separation

A perceptual and a sensorimotor component were extracted from the multivariate signal from separate experimental conditions (see *Figure 7.1*). This represented the *attunement* phase of our approach, where spatial filters were designed to target the perceptual and sensorimotor components maximally entrained to the metronome. We hereby provide the results of the qualitative assessment of the GED according to three criteria: (a) spatial activation pattern, (b) spectral profile, and (c) eigenspectrum. These results are shown and described in detail in *Figure 7.2*.

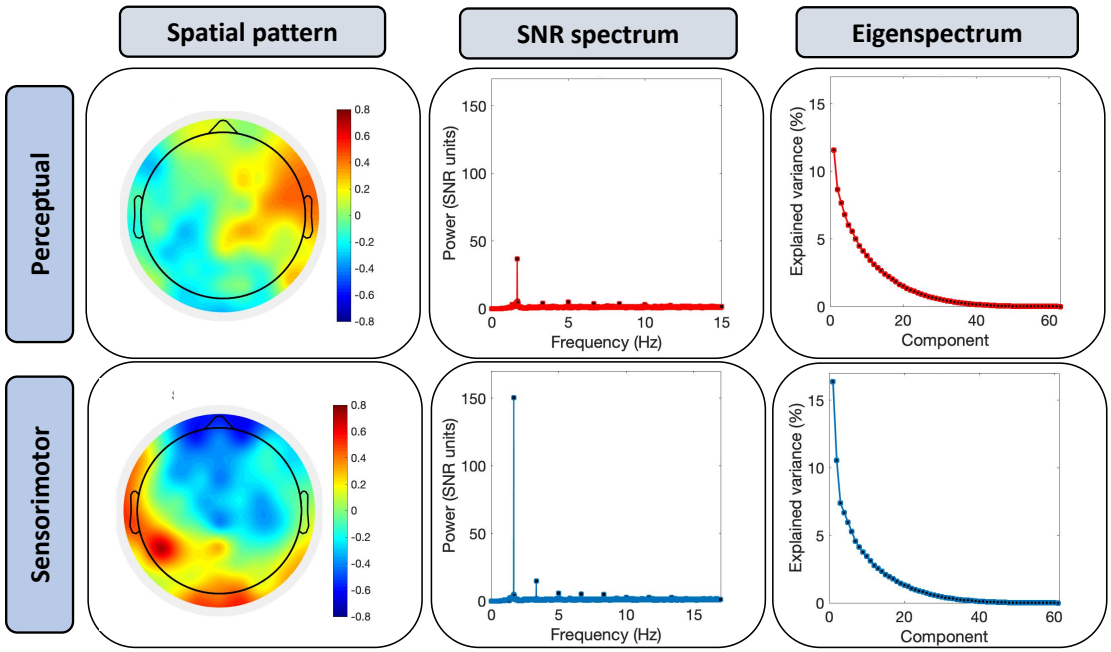


Figure 7. 2. GED source separation. The criteria for the assessment of source separation via GED are here shown at the group level (N = 19). The *spatial pattern* of the perceptual component did not exhibit the centro-frontal activation expected from EEG auditory evoked responses (Nozaradan et al., 2011, 2012, 2015), whereas the activation of the sensorimotor component exhibited both frontal negativity and a peak of activation over left centroparietal regions as previously reported during finger-tapping task performed with the contralateral hand (Rosso, Leman, et al., 2021). The *SNR spectrum* of both components was dominated by a prominent peak at the stimulation frequency (i.e., 1.67 Hz) and harmonics, indicating that our filter successfully separated from the broad-band signal a component oscillating around the center frequency of the metronome. Spectra were computed over the whole finger-tapping session, and normalized to signal-to-noise ratio (SNR) units with respect to the neighboring bins to remove the 1/f component. The *eigenspectrum* showed that the weights associated to the largest eigenvalue explained considerably more variance than the others, hence was chosen as spatial filter to for our sensor data. We observed a general additive effect of the sensorimotor component on all 3 criteria: overt movement resulted in stronger spatial activation, finer attunement to the stimulation frequency and more explained variance in the reconstructed sources. Taken together, these criteria support that the source separation successfully extracted two components attuned to the metronome's frequency during listening (perceptual component) and finger-tapping (sensorimotor component), with the latter being more effectively separated from the broadband signal. The dynamic frequency adjustment of the two components in response to perturbations was further operationalized as event-related frequency adjustment (ERFA).

2. Tempo-change perturbations

The ERFA response curves of the tempo-change conditions were modelled with 2nd order orthogonal polynomials, as the timeseries followed parabolic growth-and-decay, tracking the direction of the perturbation (see *Figure 7.3A-B-C*). This parabolic term was capable of capturing the most prominent effect in the model. Fixed effects of Direction (Negative, Null, Positive) and Component (Perceptual, Sensorimotor) on the polynomial terms were tested. The Null level (i.e., perturbation-free windows) was treated as a baseline for contrasting Negative and Positive levels of Direction, while the Perceptual level was treated as a baseline for contrasting the Sensorimotor level of Component. Model parameters were estimated with respect to the levels defined as baseline. The same polynomial model was fitted to the behavioral ERFA curves, with the exception that the Component factor was removed. The specifics of our statistical modelling are explained in detail in the *Materials and methods* section.

2.1. Neural ERFA

The orthogonal polynomial model revealed a significant main effect of Positive Direction (Estimate = 1.689, SE = 0.450, $p < 0.001$) and a significant two-way interaction between Positive Direction and the quadratic term of Time (Estimate = -8.049, SE = 4.039, $p = 0.046$). The former indicates that both Perceptual and Sensorimotor components oscillated on average significantly faster within Positive perturbated windows, the latter indicates that the higher average was accompanied by more parabolic modulation of the ERFA (i.e., inverse U-shape). Whilst the main effect of Negative Direction reached significance (Estimate = 1.160, SE = 0.450, $p = 0.010$), no significant two-way interaction with the polynomial terms was found for Negative Direction. Taken together, these effects suggest that oscillations are generally biased to follow positive frequency changes, as quantified by the ERFAs of both perceptual and sensorimotor components.

The nature of the Component becomes relevant for Negative Direction, as indicated by the significant two-way interaction (Estimate = 1.461, SE = 0.636, $p = 0.022$): the oscillation was on average significantly slower within Negative perturbated windows, but only for the Sensorimotor type. In contrast, in response to Positive perturbations, both Component types exhibited significantly faster oscillations.

A significant three-way interaction was found between quadratic term of Time, Component and Direction (Positive: Estimate = -19.129, SE = 5.711, $p < 0.001$; Negative: Estimate = -20.266, SE = 5.711, $p < 0.001$). These effects indicate that the Sensorimotor Component significantly boosted the parabolic modulation of the ERFA as compared to

the Perceptual Component, regardless of whether the metronome was speeding up or slowing down. This additive effect can be seen in *Figure 7.3A* and *B*. Results from the statistical model are reported in *Table 7.1*.

2.2. Behavioral ERFA

The orthogonal polynomial model revealed significant main effects of both Positive (Estimate = 8.816, SE = 0.567, $p < 0.001$) and Negative (Estimate = 9.276, SE = 0.567, $p < 0.001$) Directions, indicating that both signs resulted in an average frequency offset in the expected direction. The dynamic is captured by the significant interaction of the quadratic term and Direction (Positive: Estimate = -92.034, SE = 4.562, $p < 0.001$; Negative: Estimate = -85.739, SE = 4.562, $p < 0.001$), in line with the patterns of the neural ERFAs. When testing for the interaction between Direction and the linear term of the model, which indicates the position of the vertex of the parabolic curve (see(Rosso, Maes, et al., 2021), for application and interpretation of orthogonal polynomials fitted to asymmetric curves), we did not find an asymmetry across levels of Direction at our significance level alpha = 0.05, only a trend (Estimate = 9.293, SE = 4.971, $p = 0.061$). The behavioral ERFA to tempo-changes is represented in *Figure 7.3C*.

No significant difference was found at post-hoc comparisons when testing for Positive against Negative directions, neither for the main effect nor for the interaction on the quadratic terms. Results from the statistical model are reported in *Table 7.2.* Note that, since the Null level was set as a reference for both Positive and Negative in the model, effects of Direction are reported pairwise.

N = 19				
Predictors	*Estimate*	*SE*	*t value*	*p*
(Intercept)	-0.441	0.318	-1.388	0.165
Time	-0.814	2.665	-0.305	0.760
$Time^2$	0.011	2.978	0.003	0.996
Component (SM)	0.475	0.450	1.057	0.291
Direction (-)	1.160*	0.450	2.580	0.010
Direction (+)	1.689***	0.450	3.754	< 0.001
Time:Component	1.407	3.623	0.388	0.698
$Time^2$:Component	1.540	4.038	0.381	0.703
Time:Direction (-)	-1.255	3.623	-0.346	0.729
Time:Direction (+)	-1.936	3.623	-0.534	0.593
$Time^2$:Direction (-)	-6.141	4.039	-1.520	0.128
$Time^2$:Direction (+)	-8.049*	4.039	-1.993	0.046
Component:Direction (-)	1.461*	0.636	2.296	0.022
Component:Direction (+)	0.743	0.636	1.167	0.243
Time:Comp:Dir (-)	-3.250	5.124	-0.634	0.526
Time:Comp:Dir (+)	1.413	5.124	0.275	0.783
$Time^2$:Comp:Dir (-)	-20.266***	5.711	-3.548	< 0.001
$Time^2$:Comp:Dir (+)	-19.129***	5.711	-3.349	< 0.001
				** p < 0.05* *** p < 0.01* **** p < 0.001*

Table 7. 1. Neural ERFA: tempo-change.

N = 19				
Predictors	*Estimate*	*SE*	*t value*	*p*
(Intercept)	-0.709	0.560	-1.267	0.205
Time	-1.018	4.834	-0.211	0.833
$Time^2$	3.957	3.871	1.022	0.307
Direction (-)	9.276***	0.567	16.345	**< 0.001**
Direction (+)	8.816***	0.567	15.534	**< 0.001**
Time:Direction (-)	9.293	4.971	1.869	**0.061**
Time:Direction (+)	0.830	4.971	0.167	0.867
$Time^2$:Direction (-)	-85.739***	4.562	-18.794	**< 0.001**
$Time^2$:Direction (+)	-92.034***	4.562	-20.174	**< 0.001**
				** p < 0.05* *** p < 0.01* **** p < 0.001*

Table 7. 2. Behavioral ERFA: tempo-change.

3. Phase-shift perturbations

The signed area (i.e., discrete integrals) under the ERFA curves was computed as response variable, and a 3x2 factorial model was fitted: Direction (Negative, Null, Positive) x Component (Perceptual, Sensorimotor). In this context, integrals provided a measure of entity of the shift in instantaneous frequency regardless of the curve shape, while being sensitive to the direction of the shift. This sensitivity is due to the fact that the area is signed, returning positive or negative values for portions above or below the x-axis, respectively. In addition, for behavioral ERFAs, reaction times were computed as the intersection of the curve's inflection point with the time x-axis. The specifics of our statistical modelling are explained in detail in the *Materials and methods* section.

3.1. Neural ERFA

We found a significant two-way interaction between Component and Positive Direction (Estimate = 1763.756, SE = 781.745, $p < 0.024$), indicating that phase-shifts in such direction elicited significant ERFAs only when the component was Sensorimotor. The difference between components can be seen in *Figure 7.3D* and *E*, where a flat grand-average ERFA is observed for the Perceptual component due to the absence of systematic changes in instantaneous frequency across participants. It is noteworthy that the ERFA of the Sensorimotor component in Negative Direction did not mirror its Positive counterpart: rather than dropping below the baseline frequency before re-stabilizing, the curve underwent a later transition and stabilization at a higher frequency (*Figure 7.3E*). This resulted in a very small area under the ERFA in the defined post-stimulus window, and a non-significant effect in our model. Results from the statistical model are reported in *Table 7.3*.

3.2. Behavioral ERFA

A significant main effect of Direction respect to the Null level (Positive: Estimate = 9197.843, SE = 2506.882, $p < 0.001$; Negative: Estimate = 10920.459, SE = 2506.882, $p < 0.001$) was found. A post-hoc comparison between Positive and Negative levels did not yield a significant effect of Direction. This indicates that there were no differences in magnitude of error adjustments across perturbation directions. However, a t-test revealed a significant difference between reaction times to Positive and Negative perturbations (Estimate = 49.316, SE = 11.143, $p < 0.001$), indicating that participants corrected the errors significantly faster in response to Negative phase-shifts. The behavioral ERFA to phase-shifts is represented in *Figure 7.3C*. Results from the statistical model are reported in *Table 7.4* and *7.5*.

N = 19				
Predictors	*Estimate*	*SE*	*t value*	*p*
(Intercept)	557.929	413.656	1.349	0.177
Component (SM)	-617.651	552.777	-1.117	0.264
Direction (-)	-775.485	552.777	-1.403	0.161
Direction (+)	297.630	552.777	0.538	0.590
Component:Direction (-)	-24.904	781.745	-0.032	0.974
Component:Direction (+)	1763.756*	781.745	2.256	**0.024**
				** p < 0.05* *** p < 0.01* **** p < 0.001*

Table 7. 3. Neural ERFA: phase-shift.

N = 19				
Predictors	*Estimate*	*SE*	*t value*	*p*
(Intercept)	-372.017	1772.633	-0.210	0.834
Direction (-)	10920.459***	2506.882	4.356	**< 0.001**
Direction (+)	9197.843***	2506.882	3.670	**< 0.001**
				** p < 0.05* *** p < 0.01* **** p < 0.001*

Table 7. 4. Behavioral ERFA: phase-shift.

N = 19				
Predictors	*Estimate*	*SE*	*t value*	*p*
(Intercept)	490.789	7.879	62.290	< 0.001
Direction (+)	49.316***	11.143	4.426	**< 0.001**
				** p < 0.05* *** p < 0.01* **** p < 0.001*

Table 7. 5. Reaction times: phase-shift.

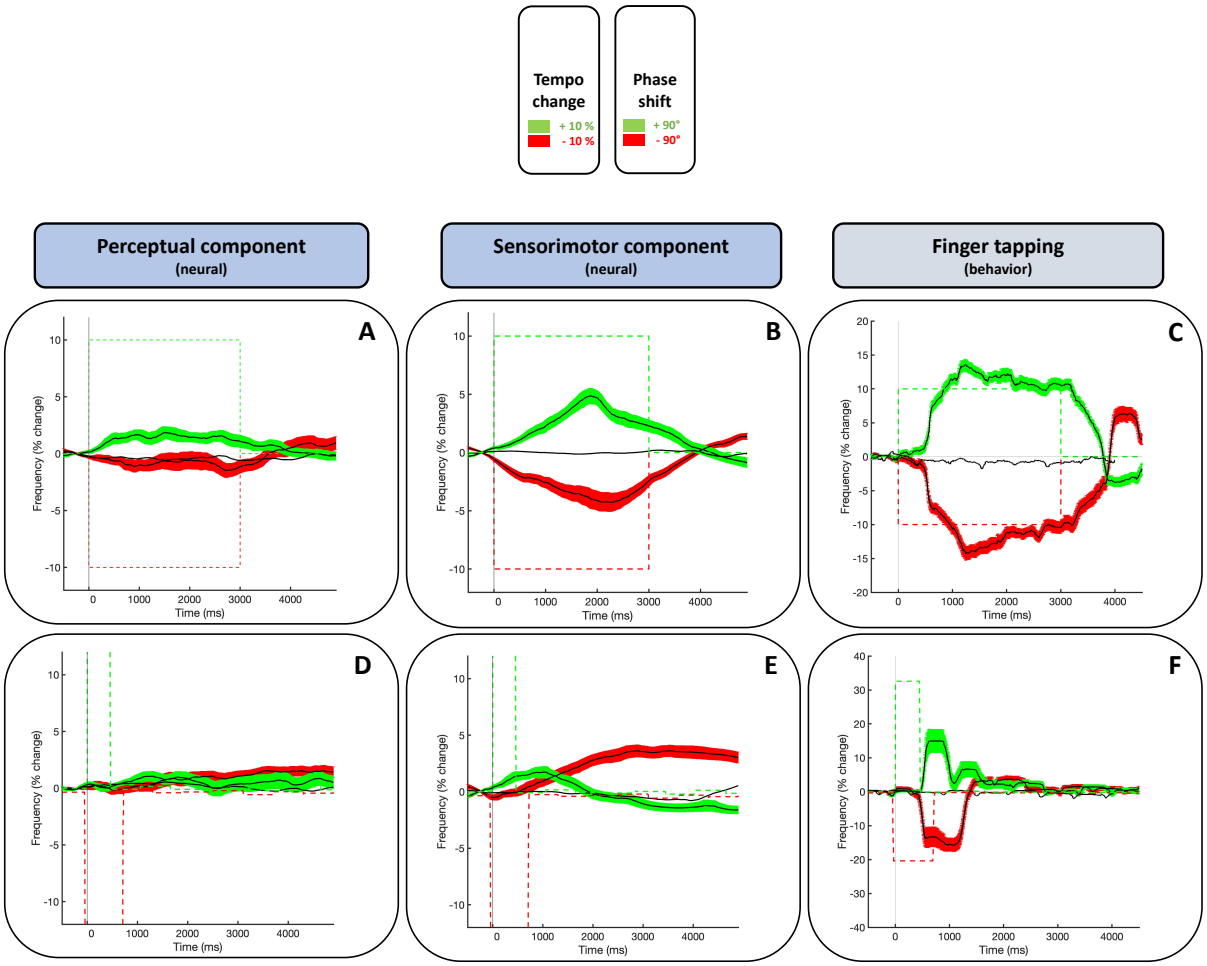


Figure 7. 3. ERFA curves. Results are here represented as grand-averages across ERFA types and perturbation types (N = 19; 19 trials per participant for every perturbation type), and expressed in percentage change with respect to the pre-perturbation stimulation frequency. Positive and negative perturbations are color-coded in green and red, respectively. For more visual clarity, the average curves are always represented as a black continuous line, whereas the contour representing the standard error of the mean (SEM) is color-coded according to the direction of the perturbation. The black continuous line without contour, approximately flat, represents the average frequency change in a sample of 19 non-perturbated time windows, where the frequency of the metronomes was stable. Dashed lines represent the instantaneous frequency timeseries of the metronome. Due to different magnitude of neural and behavioral responses, these are shown on different scales to better visualize the dynamic. Labels correspond to the following: **A- neural ERFA (perceptual component, tempo-change)**; **B- neural ERFA (sensorimotor component, tempo-change)**; **C- behavioral ERFA (tempo-change)**; **D- neural ERFA (perceptual component, phase-shift)**; **E- neural ERFA (sensorimotor component, phase-shift)**; **F- behavioral ERFA (phase-shift)**.

Discussion

The aim of the present work was to present a paradigm and a measure capable of quantifying neural entrainment from an electrophysiological brain signal, rigorously informed by the fundamental definition of the process (Lakatos et al., 2019; Novembre & Iannetti, 2018; Rajendran & Schnupp, 2019). To tackle this methodological challenge (Haegens & Zion Golumbic, 2018), we moved away from a frequency-domain representation of oscillatory components in the EEG signal. Instead, we modelled their frequency adjustment as a function of time, provided a controlled manipulation of the stimulus dynamics. In the context of a finger-tapping synchronization task to perturbated auditory metronomes, event-related frequency adjustments (ERFAs) revealed how oscillatory components entrain to the stimulus by speeding up and slowing down, tracking dynamic rhythmic changes within critical time windows.

Crucially, our experimental design allowed us to disentangle in the brain signal a perceptual and a sensorimotor component, separately attuned to the auditory metronome via GED (Michael X. Cohen, 2022), and to statistically compare their ERFAs (see *Figure 7.1*). The results showed that sensorimotor processing is critical for neural entrainment, if not a necessary condition for it to take place. When compared to perceptual ERFAs, sensorimotor ERFAs exhibited a significantly stronger modulation, congruent with the direction of the perturbation (*Figure 7.*3). It should be noted that, whilst with tempo-changes the effect was observed when either speeding up or slowing down the stimulus by 10% of its frequency, only phase-shifts of 90° in the positive direction (perceived as a shorter inter-beat interval) were tracked following the stimulus dynamics (*Figure 7.3B and E*). These patterns resemble a more smoothed version of the behavioral ERFAs computed from finger-tapping data (*Figure 7.3C and F*). On the other hand, negative phase-shifts (perceived as a longer inter-beat interval) elicited a sensorimotor ERFA qualitatively different than expected, deviating from both stimulus and behavioral dynamics. Specifically, we observed in this case an initial destabilization followed by a gradual increase in frequency, suggestive of a dissociation between the observed behavioral adjustments and their underlying neural mechanisms (*Figure 7.3E and F*).

Although in EEG data it is not possible to fully disentangle the entrainment of endogenous oscillations from regularly evoked responses which are time-locked to the stimulus, we highlight how our approach contributes to tackle this hard problem in the investigation of neural entrainment (Haegens & Zion Golumbic, 2018). We propose that identifying and statistically contrasting the perceptual and sensorimotor components serve as an initial progression towards better understanding the interplay of the mechanisms involved. Noteworthy, whilst the perceptual ERFAs can in principle be driven by changes in the

stimulation rate, the additional modulation observed in the sensorimotor ERFAs cannot be explained by a bottom-up processing alone. This is because the only difference across the two components is the engagement of the motor system, while the stimulation remains constant. Furthermore, our results reveal some features in the ERFAs which are not fully explainable by changes in the physical stimuli, calling into question the functioning of endogenous sensory rhythms in the brain and their own intrinsic dynamics. To elaborate, we reported asymmetries across positive and negative directions for both perceptual and sensorimotor ERFAs. For tempo-changes, the perceptual component exhibited a moderate frequency adjustment expected for the positive tempo change, but not for the negative tempo change (*Figure 7.3A*). If the perceptual ERFA were entirely driven bottom-up, we would expect positive and negative responses to mirror each other, because the absolute magnitude of the tempo-change was the same in both directions. However, the involvement of sensorimotor processing resulted in a significantly more prominent slowing down of the oscillation in response to negative tempo-change (compare the depth of the red curves in *Figures 7.3A* and *B*). Taken together, we argue that this evidence points to a general bias towards speeding up, which suggests the presence of intrinsic oscillatory dynamics at play. Following this argument, we propose that adjusting the frequency to the slowing metronome requires an extra deployment of neural resources, which are recruited by engaging the motor system to a greater extent (Kliger Amrani & Zion Golumbic, 2022) and therefore reflected in the enhanced sensorimotor ERFA.

We put forward two putative mechanisms to explain the interaction between perceptual and sensorimotor components. The first possibility (A) is that auditory rhythms do not entrain at all, and instead pure evoked responses are generated in sensory areas. The function of these periodic responses may be to form an amplitude-based representation of the external rhythm at the cortical level, so that endogenous motor rhythms (Morillon & Baillet, 2017; Morillon et al., 2015, 2014) can entrain by dynamically adapting their frequency via phase-based alignment. Alternatively (B), endogenous auditory rhythms might be effectively entraining to the metronome (Lakatos et al., 2016; Lakatos, Chen, O'Connell, Mills, & Schroeder, 2007; Lakatos et al., 2008, 2013, 2009), and in turn drive the alignment of endogenous motor rhythms. We argue that the evidence hereby presented leans in favor of the latter mechanism, which is also more coherent with recent neurophysiological models proposing that neural firing rates encode an abstract representation of stimulus and movement cycles', enabling both beat perception and overt synchronization to the beat (Cannon & Patel, 2021). Backed by evidence from monkeys (Cadena-Valencia, García-Garibay, Merchant, Jazayeri, & de Lafuente, 2018; Gámez, Mendoza, Prado, Betancourt, & Merchant, 2019) and humans (Bengtsson et al.,

2009; Chen, Penhune, & Zatorre, 2008; Grahn & Brett, 2007; Teki, Grube, Kumar, & Griffiths, 2011), the authors proposed that the supplementary motor area (SMA) is the key structure responsible for this cyclic sensorimotor process, working as interface between auditory pathways and motor areas.

For the sake of completeness, we provide as *Supplementary material 7.B* the grand-average activation timeseries for perceptual and sensorimotor components during tempo-changes, to give a visual impression of the evoked responses in the signal from which ERFAs were computed. Although our task was not designed for ERPs, it is still noteworthy to see that there is no visible pattern of evoked responses expected by the stimulation rates. Additionally, we compared in a series of simulations ERFAs produced by an oscillatory model and ERFAs produced by an alternative model of evoked responses (see *Supplementary material 7.C*). Overall, the ERFAs computed from oscillations appear to be more robust to varying levels of noise and are still reliable in conditions of poor signal-to-noise ratio, which better approximates the reality of signals recorded with EEG. Given all the above, we argue that the parabolic ERFAs observed in *Figures* 7.*3A* and *B* are better explained by an oscillatory model, as compared to evoked responses passively tracking changes in stimulation rate.

For phase-shifts, the asymmetry across directions was more radical: in response to negative perturbations, the sensorimotor ERFA showed a gradual transition towards higher frequency (*Figure 7.3E*). Given this class of perturbation was more localized in time and more prominent than a sustained -10% tempo-change, it may be less demanding for the brain to gradually speed up an entrained oscillation and catch up with the beat over some cycles, rather than suddenly slowing down to track the dynamic of a phase-shift. Evidence from neuropathology supports the idea that a bias for faster tempi is a feature in healthy adult population, which can be impaired by lesions in the cerebellum. When such structure, critical for event-based timing (Schwartze, Keller, & Kotz, 2016), is compromised, high stimulation rate becomes detrimental for the neural tracking of the beat as quantified by frequency tagging (Nozaradan et al., 2017), while a case study suggests that behavioral synchronization improves when stimulation rate is below the spontaneous rate of movement (Moumdjian et al., 2022). Another observation in favor of the entrainment of sensory rhythms is the fact that, whilst perceptual ERFAs to tempo-changes exhibit a certain degree of curvature (*Figure 7.3A*), they are flat in response to phase-shifts (*Figure 7.3D*). Endogenous oscillations would in fact need several cycles to entrain, a condition that is met with sustained tempo-changes.

Finally, the topography of the perceptual component (*Figure 7.2*) does not resemble the fronto-central cluster expected from auditory evoked responses (Nozaradan et al., 2011,

2012, 2015). We observed instead a more distributed pattern, whose activation was considerably weak compared to the sensorimotor component. Active motor engagement resulted once again in an additive effect, with maximal left centro-parietal activation as previously reported during finger-tapping performed with the right hand (Rosso, Leman, et al., 2021). Although, in line with recent research (Cheng, Creel, & Iversen, 2022; Kliger Amrani & Zion Golumbic, 2022), the evidence hereby provided points at a special influence of the motor system on neural entrainment, we cannot rule out that the differences observed across components may be partly explained by improved signal-to-noise ratio in the sensorimotor component. Furthermore, we do not know whether the same neural adjustments would have been observed in the absence of the motor requirement of the finger-tapping task, and to what degree. We emphasize the need for future research to expand on the present experimental design and better address the role of motor involvement in neural entrainment.

Alongside the neural ERFAs discussed so far, we also analyzed behavioral ERFAs as the change in the instantaneous frequency of finger-tapping during overt error-correction responses to perturbations. We observed highly consistent dynamics at the group level, but also some degree of interindividual variability in response to phase-shifts, showing that some participants adopted different strategies to correct the synchronization error by deviating from the stimulus dynamic. The systematic classification of such strategies is out of the scope of this paper. Adaptation to tempo-changes accurately tracked the stimulus dynamics, and results were consistent with the reported neural responses: a significant difference in average frequency and parabolic curvature was found in both positive and negative perturbations as compared to the baseline. When testing the two directions at post-hoc comparisons, no significant differences were found in the dynamics of the behavioral responses. As for the adaptation to phase-shifts, participants adapted as expected in both directions. Although no significant difference in the entity of the correction across directions was found, a test on the reaction times reveled that participants were significantly faster in adapting to negative phase-shifts (perceived as a longer inter-beat interval).

A strength of ERFA to be highlighted is that it provides an overarching analysis framework for signals of different nature. The temporal dynamics of rhythmic stimulation, behavioral responses and electrophysiological signals were all processed in terms of instantaneous frequency, allowing direct comparisons across different measurements. In the particular case of finger-tapping, data were processed modelling rhythmic behavior as an oscillator (Ole Adrian Heggli et al., 2019; Rosso, Maes, et al., 2021), congruently with a neural oscillatory framework. ERFAs were also comparable across perturbation types, namely tempo-changes and phase-shifts. Despite the discussion on different cognitive

mechanisms underlying phase and tempo corrections (B. H. Repp, 2001b; Bruno H. Repp & Keller, 2004), from a signal processing perspective they can be operationalized as the same phenomenon on different timescales, expressed in the same unit of measure. Instantaneous frequency was in fact computed entirely based on the rate of change of an oscillation's phase, and its expression in Hz units is just a matter of re-scaling (Michael X. Cohen, 2014). Phase information is needed for estimating changes in frequency, and is a necessary condition for operationalizing entrainment according to its fundamental definition (Rajendran & Schnupp, 2019).

Our work was mainly focused on oscillatory dynamics recorded with EEG. Despite the constraints of the poor spatial resolution and low signal-to-noise ratio characterizing the technique, we address the importance of applying our experimental paradigm and metric in populations affected by neurological deficits, to generate testable predictions on the functional role of neuroanatomical structures compromised by the pathology (Cannon & Patel, 2021). Furthermore, future research may deploy the paradigm with magnetoencephalography or intracranial recordings, to investigate entrained activity in cortical and subcortical anatomical structure during sensorimotor synchronization.

Conclusions

The major methodological contribution of our work consisted of a paradigm and a measure for investigating neural entrainment in human participants, optimized for non-invasive electrophysiological recordings. By perturbating isochronous auditory metronomes in tempo and phase during a finger-tapping task, we induced behavioral synchronization errors and showed that oscillatory neural components dynamically adjusted their frequency to stimulus changes during error-correction responses. By means of spatial filters design, we were able to disentangle perceptual and sensorimotor oscillatory components from the multivariate EEG signal, revealing that active engagement of the motor system enhanced neural entrainment. This evidence, along with clues of intrinsic brain dynamics not explicable by bottom-up processing of the stimuli, strongly suggests that actual neural entrainment underlies tracking and sensorimotor synchronization to dynamic auditory rhythms. In addition to these fundamental findings, ERFA proved to be a sensitive measure of neural entrainment, reflecting an oscillatory model of brain functioning while mitigating the influence of bottom-up evoked responses.

Author Contributions

Mattia Rosso: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing – Original Draft, Visualization, Project administration, Funding acquisition. **Bart Moens**: Software, Resources. **Marc Leman**: Writing – Review and Editing, Supervision. **Lousin Moumdjian**: Conceptualization, Writing – Review and Editing, Project administration.

Competing Interest Statement

The authors declare no competing interests.

Acknowledgments

The present study was funded by Bijzonder Onderzoeksfonds (BOF) from Ghent University (Belgium), in the context of a joint-PhD project with the University of Lille (France) (I-SITE ULNE program; grant number: 01D21819). The authors are grateful to Ivan Schepers for building the hardware of the finger-tapping device, to Canan Nuran Gener for her precious help in collecting the data, and to the three anonymous reviewers whose thorough comments considerably improved the quality of the manuscript.

8

Sensorimotor synchronization in normal and pathological ageing: an analysis toolkit

Mattia Rosso [A,B], Andres von Schnehen [B], Lise Hobeika [B], Marc Leman[A], Séverine Samson [B]

[A.] IPEM - Institute for Systematic Musicology; Ghent University; Ghent, Flanders, 9000; Belgium.

[B.] PSITEC - Psychologie: Interactions, Temps, Emotions, Cognition - ULR 4072; University of Lille; Lille, Hauts-de-France, 59650; France.

Introduction

This Chapter represents a corpus of methodological contributions in the context of an ongoing collaboration between *IPEM - Institute for systematic musicology* (Ghent University, Belgium) and *PSITEC Laboratory - Psychologie: Interactions, Temps, Emotions, Cognition* (University of Lille, France). Historically, the joint effort of the two research groups has been focused on investigating the potential of music-based interventions for elderly people affected by neuro-cognitive disorders (NCDs) (Desmet, Lesaffre, Six, Ehrlé, & Samson, 2017; Ghilain et al., 2020; Ghilain, Schiaratura, Singh, Lesaffre, & Samson, 2019; Hobeika et al., 2021; M. Lesaffre, Moens, & Desmet, 2017). This diagnosis was introduced in the DSM-5 to refer to different classes of dementia, including Alzheimer Disease. Moreover, NCDs are classified as 'major' or 'mild' based on the severity of the cognitive decline in respect to a level of performance previous to the onset of the pathology.

What is it about music that makes it a potential tool in the rehabilitation of cognitive and social functions in elderly people with NCDs? To this date, research has left open questions as to whether these provide specific benefits as compared to non-pharmacological interventions consisting of other pleasant activities (Baird & Samson, 2015; Palisson et al., 2015; Samson, Clément, Narme, Schiaratura, & Ehrlé, 2015). The case for music-based interventions is made based on the sensitivity and emotional responsiveness of patients with dementia to music (Sihvonen et al., 2017; J. T. van der Steen et al., 2018), with larger benefits when they are actively engaged in singing and/or moving along with a musical beat (Sakamoto, Ando, & Tsutou, 2013). This clinical scenario fits well within the framework of the dissertation, where the social component of a human interaction is brought in by the physical presence of a music therapist, and mediated by the rhythms of musical stimuli and bodily motion.

This sort of musical tasks leverages on the relatively spared sensorimotor synchronization (SMS) to musical rhythms up to a certain stage of disease progression (Ferreri et al., 2016; Goldman, Baty, Buckles, Sahrmann, & Morris, 1999). Crucially, when tapping along to either metronomes or familiar songs, no significant differences in SMS accuracy and consistency are found across cognitively impaired patients and age-matched controls (Ghilain et al., 2020). However, the differences emerge when accounting for the severity of the disease. Patients categorized as affected by major neurocognitive disorder (Major NCD) perform less accurate and less consistent finger-taps when synchronizing to metronomes, in comparison with Minor NCD and No NCD groups, and consistency decreases with higher MMSE scores (Hobeika et al., 2022). When tapping to music, and in physical presence of a musician tapping along with the patient, the performance of this

group is comparable to the No NCD group. These studies highlight that the presence of a human model, be it physical or tele-mediated (Ghilain et al., 2020), provides a visual support that is crucial for a better execution of a rhythmic task. For advanced stages of the disease, the physical presence becomes more important (Hobeika et al., 2022). In dynamical terms, the informational basis of the interaction (R. C. Schmidt et al., 1990) is maximized by combining visual and auditory coupling, to take advantage of their specific properties (see Chapters 2-4).

SMS is thought to be the core function driving interactive music-based interventions, as its stimulation is foundational to their effectiveness (see (von Schnehen, Hobeika, Huvent-Grelle, & Samson, 2022), for an updated review on the topic). To thoroughly comprehend its role, it's vital to examine how motor, expressive, and empathic processes connect to this basic function during a rhythmic interaction. This is because activating brain networks associated with movement can modulate emotion and stimulate cognition (Zatorre, Chen, & Penhune, 2007). Major NCD patients demonstrate reduced motor and socio-emotional engagement, as evidenced by diminished overall body motion and emotional facial expression, as well as gaze direction during the finger-tapping task (Hobeika et al., 2021). The findings suggest that these measures could be indicative markers of disease progression. For a comprehensive understanding of how patients with neurocognitive impairments respond to music at various levels, it is crucial to employ multimodal non-invasive monitoring technology within the framework of musical interaction (M. Lesaffre et al., 2017). Comprehending how disease progression impacts the mechanisms underpinning music interactions is crucial for informing and tailoring music-based rehabilitations. However, years of experience have revealed significant limitations that must be considered when designing experiments involving this particularly vulnerable population (Hobeika et al., 2022). To elucidate further, some of these limitations include:

- The application of sensors or measurement devices, which are generally well-tolerated by young healthy adults, can be perceived as overly invasive by individuals with NCDs, potentially inducing stress and fear responses (Desmet et al., 2017). The impracticality of using electrophysiological measurements like EEG, due to their long preparation time, perceived invasiveness, and physical discomfort, limits the feasibility of exploring neural dynamics using such tools. This constraint necessitates a greater focus on optimizing behavioral measurements within a comfortable and familiar setup.

- The tolerance of these patients for long experimental sessions is notably limited, especially when the procedure is prolonged, repetitive, and requires

continuous movement. Tolerance and compliance are further influenced by age progression, disease severity, and associated cognitive capacity and mobility. This imposes strict limitations on the amount of data collected, which challenges the generalizability of the findings and the reliability of techniques susceptible to low signal-to-noise ratio (SNR) and high behavioral variability.

- In terms of auditory stimuli selection, a bias towards pleasant and familiar ones is often required to ensure the task remains comfortable, enjoyable, and capable of maintaining a suitable level of sustained attention. Although metronomes conveniently control for preference, familiarity, and complexity in the stimulus, they fall obviously on the other extreme end of this spectrum.

- Lastly, maintaining participant engagement throughout the task necessitates the active presence of an experimenter, a therapist, a music teacher, or some form of human guide. While this setup may be convenient when explicitly studying collaborative or imitative behaviors, it can be problematic when requiring unguided control conditions or spontaneous, uninstructed synchronization. While the combination of auditory and visual couplings may be beneficial for rehabilitation, it is not possible to manipulate them independently when a human guide has to be present, making it challenging to discern the individual contributions of sensory inputs to the dynamics of the interaction.

In summary, experimental procedures for this population must prioritize non-invasiveness, brevity, minimal repetition, engagement, human support in a familiar environment, and a large number of participants to compensate for the limited data collected in individual sessions.

It becomes evident that paradigms designed for young, healthy adults, such as *drifting metronomes* for dyadic entrainment (Chapters 2-4) and *event-related frequency adjustment* (*ERFA*) for error-correction in SMS (Chapter 7), are not feasible in their original forms. Moreover, the necessary adaptations could significantly alter these paradigms, compromising their intended purposes or yielding insufficient data at the expense of the analyses. In balancing methodological precision and feasibility, research involving this population must prioritize the latter. PSITEC has successfully achieved this by collecting substantial datasets from patients across French healthcare structures in the intervals between clinical sessions, day hospitals, or neuropsychological assessments. Consequently, this collaboration is focused on working with extensive pre-collected

datasets, guided by the methodological principles presented in this dissertation. Moving from these limitations, this Chapter aims to maximize inference through data analysis, developing approaches centered on the behavioral dynamics of rhythmic interactions involving patients with NCDs.

Methodological contribution

The general experimental design of the studies conducted at PSITEC involves a mixed factorial design where the Group is treated as a between-subjects factor, with levels referring to the severity of neurocognitive disorders (No NCD, Minor NCD, Major NCD). Complementarily, neuropsychological tests such as the mini mental state examination (MMSE) (Tombaugh & McIntyre, 1992) enable the treatment of disease severity as continuous predictor, distinguishing its contribution from the component of impairment attributable solely to aging (Hobeika et al., 2022). Stimuli (Metronomes, Music) and Social (Live, video) factors are manipulated within-subject. In keeping with the experimental design of the collected datasets, the contribution of this Chapter lies in processing timeseries data to yield new outcome measures.

The same 3x2x2 mixed-ANOVA models based on the factorial structure (Group x Stimuli x Social) will be fitted to the response variables returned by our analyses. Importantly, these new outcome measures aim to provide more comprehensive information on the dynamics of timeseries data captured from musical interactions, complementing the average-based measures used thus far. Maintaining the dual focus on social and individual dimensions of rhythmic interactions throughout this dissertation, we present two separate methodological contributions based on two datasets recorded from healthy and cognitively impaired elderly participants at Bateliers Day Hospital (Lille University Medical Center, Lille, France). After presenting the analyses, we will present the results in the form of descriptive statistics. While these data will be further subjected to the above-mentioned statistical models, such exploration lies beyond the scope of the present discussion and the author's direct influence. All analyses were implemented in Matlab 2019a (MathWorks Inc., USA).

Dataset #1 – Spatiotemporal dynamics of body-sway (Ghilain et al., 2020; Hobeika et al., 2021)

The first dataset was collected in a study that investigated the natural movement responses to music in individuals with dementia (Ghilain et al., 2020), specifically focusing on the activation of weight sensors from a force plate (Hobeika et al., 2021). The research aimed to explore the potential of music-based interventions as non-pharmacological approaches to enhance emotional, social, and cognitive functioning in patients with neurodegenerative diseases, such as Alzheimer's Disease (AD). A key objective of the study was to understand how motor involvement could increase the benefits of these interventions. Participants (N = 194) were asked to tap in synchrony with musical rhythms while a musician tapped along with them. The study compared the musician's live performance to a pre-recorded video of the performance, in terms of the impact on the participants' sensorimotor abilities. In this context, measuring full-body motion by means of the force plate provides valuable insights into one aspect of the participants' engagement in the musical interaction (Desmet et al., 2017; M. Lesaffre et al., 2017). This contributes to a comprehensive understanding of SMS and its interplay with the motor, social, and emotional components of the overall engagement.

For the most part, the methodologies presented in previous Chapters of this dissertation have focused exclusively on the temporal dynamics of human rhythmic movement and brain activity, while the spatial component was minimized and held constant across studies and experimental conditions. In particular, behavioral measures were based on finger-tapping, which involved unidimensional movement along the vertical axis. Here, the multivariate recordings provided by the four sensors, combined with information on their topographical arrangement, allowed us to introduce a spatial dimension to the behavioral data, enabling us to examine the spatiotemporal dynamics of body motion during a rhythmic interaction. In incremental levels of complexity, the analysis pipeline developed for this dataset will yield: 1) a measure of *quantity of motion* (*QoM*) based on the work of Desmet et al. (Desmet et al., 2017); 2) a set of measures from *recurrence quantification analysis* (*RQA*) (Afsar et al., 2018; Demos & Chaffin, 2017; Demos et al., 2018); 3) a set of measures from *joint recurrence quantification analysis* (*JRQA*) (Demos & Chaffin, 2017). These techniques aim to answer the following questions:

1) "How much motion is in the data?";
2) "Is there temporal structure in such motion?";
3) "How much of that structure is explained by the interaction with the musician?".

Experimental apparatus

The recording device was composed of a square wooden force plate (90 x 90 cm) situated on a frame with four calibrated weight sensors, one at each corner in a squared configuration. These delivered weight measurements and were individually read by an Arduino® Due open-source prototyping platform, which facilitated the calculation of movement direction and magnitude. A software application was created to read multiple force plates simultaneously, and to collect data from various measuring devices, such as pressure sensors, audio, and video recording. The experimental arrangement included two force plates designed by IPEM at Ghent University, each with a chair attached, one for the patient and the other for the musician. The force plates featured four sensors (one at each corner) to assess the movements of the person seated in the chair. A small table was affixed to each chair, offering a convenient position for tapping in time with the music. A microphone was positioned beneath the table to capture the tapping sounds of the participants as audio signals. Two webcams were set up to record videos of both balance boards during the experiment. In addition, a projection screen with a projector behind it was situated at the back of the experimenter's board to allow for prerecorded video projections required for the video aspects of the experiment. The setup was overall non-invasive and suited for use in an ecological setting. For a population where mobility is often heavily hindered by age and/or the progression of the disease, it provided a feasible solution for multimodal acquisition, complementing SMS as measured by finger-tapping with a capture of full-body movement from a comfortable position. The force plate system was developed aiming at providing a balance between functionality for the patient, sensitivity for measuring smaller movements, and data reliability. A more detailed technical description is provided in (M. Lesaffre et al., 2017).

Test datasets

In order to test our analysis before applying it to the patients' dataset, we simulated and recorded data from the recording device under controlled conditions. Test datasets included a set of stereotyped behaviors exhibiting clearly identifiable spatiotemporal features. We will refer to the simulated data and the data recorded by the experimenter under controlled conditions as *'simulated data*' and *'mock data*,' respectively. While both serve as placeholders for real data, providing a ground truth for each step of the process as they are fed into the analysis pipeline, each set comes with its unique advantages.

Simulated data allowed us to have simulated scenarios of ideal behavior, with the possibility to systematically and independently manipulate parameters of interest. For

instance, this enabled us to test the robustness of our analyses to different SNR levels, different levels of variance and noise in the signal across channels due to unbalanced calibration, and different levels of variability across parameters of behavior. Mock data, on the other hand, contain the empirical information on the actual SNR and calibration levels of the device, while recording stereotyped behavior from a real human being. Both placeholders were necessary to properly test our methods, accounting for both generalizability and the contingencies brought in by a particular device used in a particular experimental context.

Stereotyped behaviors.

A representative sample of behaviors was chosen based on spatiotemporal features of interest. As for the time dimension, we were interested in assessing our analyses in conditions of synchronous and asynchronous movement with respect to an isochronous auditory cue. As for the space dimension, it was crucial to assess whether results are invariant to differences in spatial patterns. The behaviors analyzed were defined as follows:

1. *Sitting (no movement)*. The participant sits without moving. This accounts for the baseline activity of the sensors, reflecting the participant's weight in the absence of movement. Simulated as uniform white noise.

2. *Random movement (asynchronous)*. The participant moves randomly, without a recognizable pattern. Simulated with pink noise added to each channel, on top of the baseline white noise. Pink noise is suitable for simulating autocorrelation structure in human movement.

3. *Bouncing (synchronous)*. The participant bounces on the chair by pressing uniformly on all sensors, synchronizing with a rhythmic auditory stimulus. Simulated as four zero-phase sine waves oscillating at the stimulus frequency, plus uniform white noise.

4. *Circular motion (synchronous rotation, clockwise)*. The participant shifts their weight from one sensor to another in a clockwise circular motion, paced by an auditory rhythmic stimulus. Simulated as four sine waves with $\pi/2$ phase offsets, oscillating at 1/4 of the stimulus frequency, plus uniform white noise.

5. *Back and forth (synchronous)*. The participant shifts their weight back and forth between the two pairs of sensors placed at the front and the back of the plate, paced by an auditory rhythmic stimulus. Simulated as two pairs of sine waves divided by the longitudinal plane, with π phase offset, oscillating at 1/2 of the stimulus frequency, plus uniform white noise.

6. *Left and right (synchronous)*. The participant shifts their weight left and right between the two pairs of sensors placed at the sides of the plate, paced by an auditory rhythmic stimulus. Simulated as two pairs of sine waves divided by the frontal plane, with π phase offset, oscillating at 1/2 of the stimulus frequency, plus uniform white noise.

7. *Left and right (asynchronous)*. The participant shifts their weight left and right between the two pairs of sensors placed at the sides of the plate, ignoring the auditory rhythmic stimulus. Simulated as two pairs of sine waves divided by the frontal plane, with π phase offset, oscillating at 1/2 of the stimulus frequency, plus uniform white noise.

8. *Back and forth / left and right (synchronous, with transition)*. The participant alternates between behaviors 5 and 6 every 10 seconds, based on a visual timer. Simulated as behaviors 5 and 6, with the pairs of sensors undergoing the phase offset changed every 10 seconds.

Simulated data were simulated over a time period of 60 seconds at a sampling rate of 500 Hz. White noise was added to the timeseries (SNR = 10). Mock data were recorded and stored as .wav files with seven aligned tracks sampled at an audio rate of 44.1 kHz, later down sampled by a factor of 88 to match the simulated timeseries. The tracks contained the following data: audio signal (metronome or music), tapping timeseries, an empty track, front-left sensor, front-right sensor, back-right sensor, and back-left sensor. Sensor activations over a period of 20 seconds are shown in *Figure 8.1* and *Figure 8.2* for simulated and mock data. Offsets on the y-axis are introduced for visualization purpose.

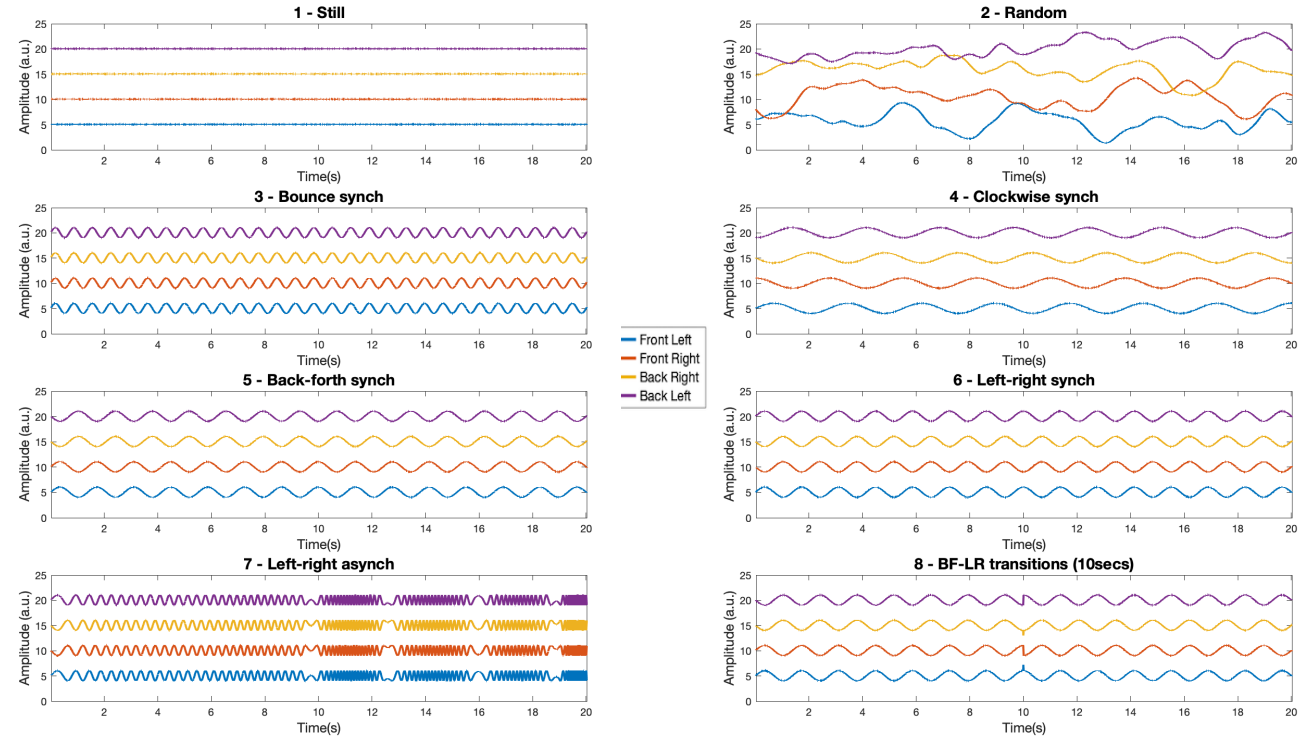


Figure 8. 1. Simulated data for stereotyped behaviors.

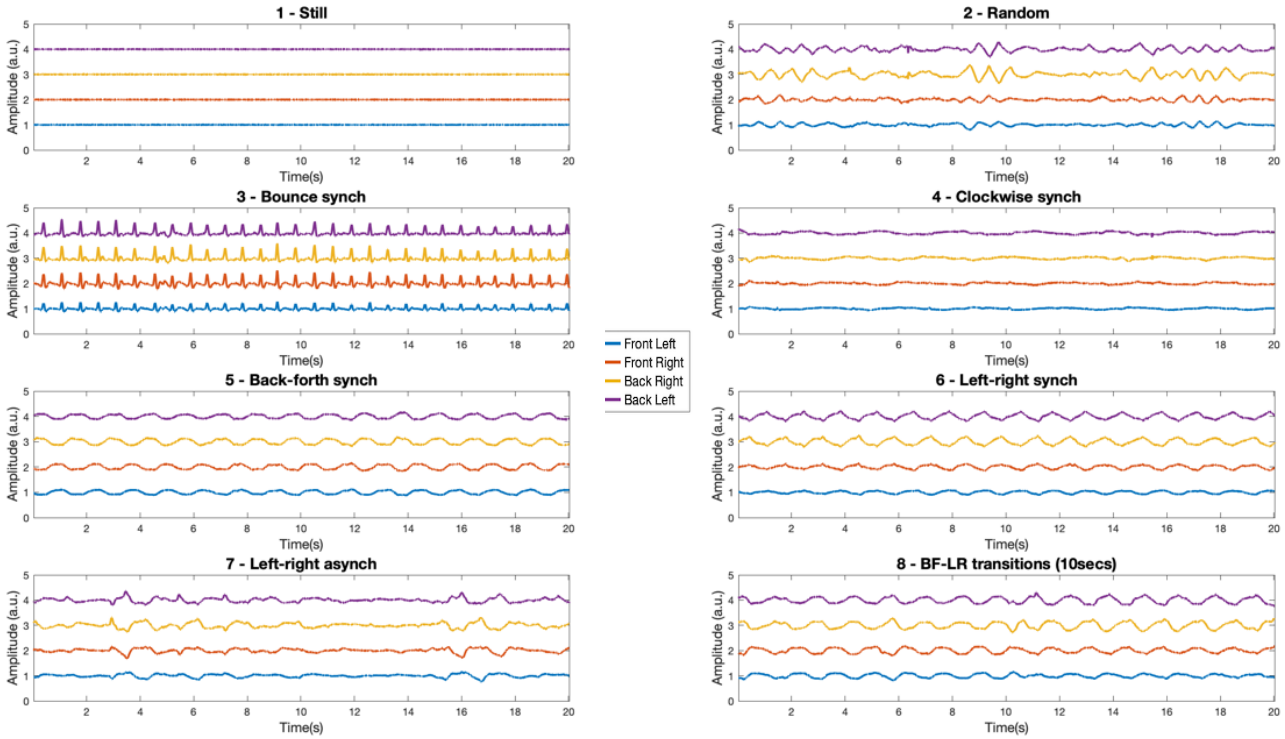


Figure 8. 2. Mock data for stereotyped behaviors.

Data cleaning via Independent Component Analysis (ICA)

Upon visual inspection of the force plate timeseries recorded in (Ghilain et al., 2020), a systematic artifact caught our attention. The artifact manifested as recurrent series of sharp transient activity, which overlaid nicely with data recorded from finger-tapping timeseries acquired separately with a dedicated sensor. Furthermore, it was evident that the artifact was systematically more prominent in the top-right sensor, due to its placement near a support stand. This represents a serious issue for our analyses, as it is reasonable to expect that differences in rate and amplitude of the finger-tapping could significantly explain differences in the QoM, while their temporal structure could be mistaken for temporal structure in body sway. For these reasons, our pipeline had to start by separating and removing the source of artifactual activity from the data.

In order to that, wed leveraged the following facts: 1) a multivariate dataset, such as our force plate, allows for source separation, which can be used for artifact removal; 2) finger-tapping timeseries were recorded by a separate device, providing a ground-truth source for the estimation of artifactual activity; 3) the artifact was empirically identified as being generated by a specific behavior, which, with respect to body sway, represents an independent source in a conceptual, statistical, and mathematical sense. Taken together, these reasons point to Independent Component Analysis (ICA) as the tool of choice for artifact removal. The method is widely adopted in EEG literature for the removal of stereotyped artifacts such as eye-blinks (optionally recorded with EOG to provide a ground-truth) (M. X. Cohen, 2014; Luck, 2014), it was adopted in the EEG studies presented in this dissertation (Chapters 5-7), and expanded upon in a separate methodological work (Vidal et al., 2021). To the best of our knowledge, this is the first application to force plate data for source separation.

To test the ICA approach, we separately recorded 60 seconds of finger-tapping behavior synchronized with a metronome on the finger-tapping sensor. We then imported this data, replicated the timeseries in a matrix of the same size as the force plate mock data, scaled to give more weight to the top-right sensor and less weight to the bottom-left sensor, and finally added them to the data to generate a mixed data matrix. The goal of our test was to un-mix the matrix to retrieve (estimate) the artifact and remove its contribution to the data, returning to the mock data as they were before the mixing. We opted to adopt the JADE algorithm (Rutledge & Jouan-Rimbaud Bouveresse, 2013). Like many others, the algorithm is based on the kurtosis operator, but unlike many others is deterministic, meaning that every iteration would result in the same independent component estimations. Moving from the assumption that signal is different from noise

due to non-normal distribution, the algorithm identifies independent components by maximizing their kurtosis.

The output of JADE is a matrix of eigenvectors (in the columns), which we multiplied by the mixed data matrix to produce independent component (IC) scores or activation timeseries. The forward model, which maps the projection of the sources onto the data in sensor space, is calculated as the inverse matrix of the eigenvectors' matrix. It is useful to visualize the topography of such model to verify that the weights map onto the sensors as expected. In our case, they should closely match the distribution of weights used to scale the artifact before mixing with the mock data (see *Figure 8.3A*). After estimating the IC scores, the artifactual component was identified as the one having the highest Pearson's correlation coefficient with the finger-tapping timeseries. To clean the data from the contribution of the artifact, we zeroed the column of weights associated with the target component in the projection (forward model) matrix, and back-projected the components onto the data via matrix multiplication. Alternatively, one may use a subset of the desired columns for performing matrix multiplication, or calculate the contribution of the artifactual component to each channel and then subtract that contribution from the original data. The operation returned the original mock data, without the contribution of the artifact (see *Figure 8.3B*). The test, successfully replicated on all mock recordings, showed that the approach is reliable for cleaning force plate data from the tapping artifact.

1. Quantity of Motion Analysis (body-sway)

Building on the methods proposed by (Desmet et al., 2017), we implemented a calculation of QoM based on the embedding of sensor data in a 2-dimensional Cartesian representation. The rationale behind this approach is to transition from a multivariate representation of the data in the time domain to a representation of movement trajectories, based on changes in the center of gravity over the squared surface of the force plate. By representing the data in terms of relative changes in activation across sensors, the data points can be interpreted as deviations from the origin of the Cartesian space or, in other words, as deviations from the center of the force plate.

Data were centered around 0 by de-meaning every channel. Next, an 'angles' matrix of the same size as the data matrix was computed, consisting of repetitions of the [0°, 90°, 180°, 270°] row. The X and Y components of the Cartesian representation of the data were computed by element-wise multiplying the data matrix with the cosine and sine of the angles, respectively. Data were further centered in Cartesian space by subtracting the means of each quadrant from all data points. Data points exceeding the overall mean by

more than 3 standard deviations were considered outliers and removed. Subsequently, Cartesian coordinates were converted to polar coordinates by computing the angle (θ) and magnitude (ρ) of each (x, y) vector. QoM was finally calculated as the average ρ, indicating the average magnitude of the 2-dimensional vectors away from the center of the space.

Simulated data are particularly useful to visually assess the validity of the approach, define the ranges of QoM expected under ideal conditions, and see how they compare across different types of motion. *Figure 8.4* shows the simulated data as represented in Cartesian and Polar coordinates, along with the associated QoM. It is interesting to note that the expected QoM for sitting and bouncing behaviors (#1 and #3) falls within a very close range, despite the apparent difference in variance, and tends to zero in the absence of noise. This demonstrates that QoM is not sensitive to the variance in sensor activation as long as participants do not move away from the center of gravity. For this reason, we can conclude that the measure is selectively quantifying body sway, rather than motion in general. Furthermore, behaviors #4, #5, #6, #7, and #8 cluster in a narrow range greater than zero, indicating that QoM is not sensitive to the direction, periodicity, or transitions of body sway.

Mock data, on the other hand, are useful for examining how variability and autocorrelations in real human motion, along with sensor calibration and mechanical coupling of the setup, can bias the distribution of data in Cartesian and polar spaces. *Figure 8.5* shows the mock data as represented in Cartesian and Polar coordinates, accompanied by the associated QoM. When comparing simulated and mock data, it becomes apparent that while the empirical patterns are consistent with the simulations, some differences emerge. The most noticeable of these are the observations that the ranges of QoM appear to depend on the direction of movement, suggesting an unbalanced calibration of the sensors, and that mechanical coupling biases the distribution of data points towards the origin of the space. For instance, examining the extreme case of circular motion (#4), it becomes evident that it is not possible to produce a hollow circle due to the spatial dependency imposing covariance on the sensors. It is important to emphasize that the patterns of stereotyped motion presented here will never be replicated by real participants, especially not by NCD patients. However, this is arguably a crucial step in gaining insight into the nuances of the measure we intend to adopt and, most importantly, to ensure that it quantifies what it is supposed to quantify.

The analysis hereby presented was applied to the recordings of 164 patients (N = 164) out of the total of 194 included in the original dataset. For 30 participants, two or more conditions were either missing or exhibited major issues in the recording, which made the

analyses unfeasible. *Figure 8.6* shows the average values of QoM in the different experimental conditions. Mean and standard deviation are reported in *Table 8.1*.

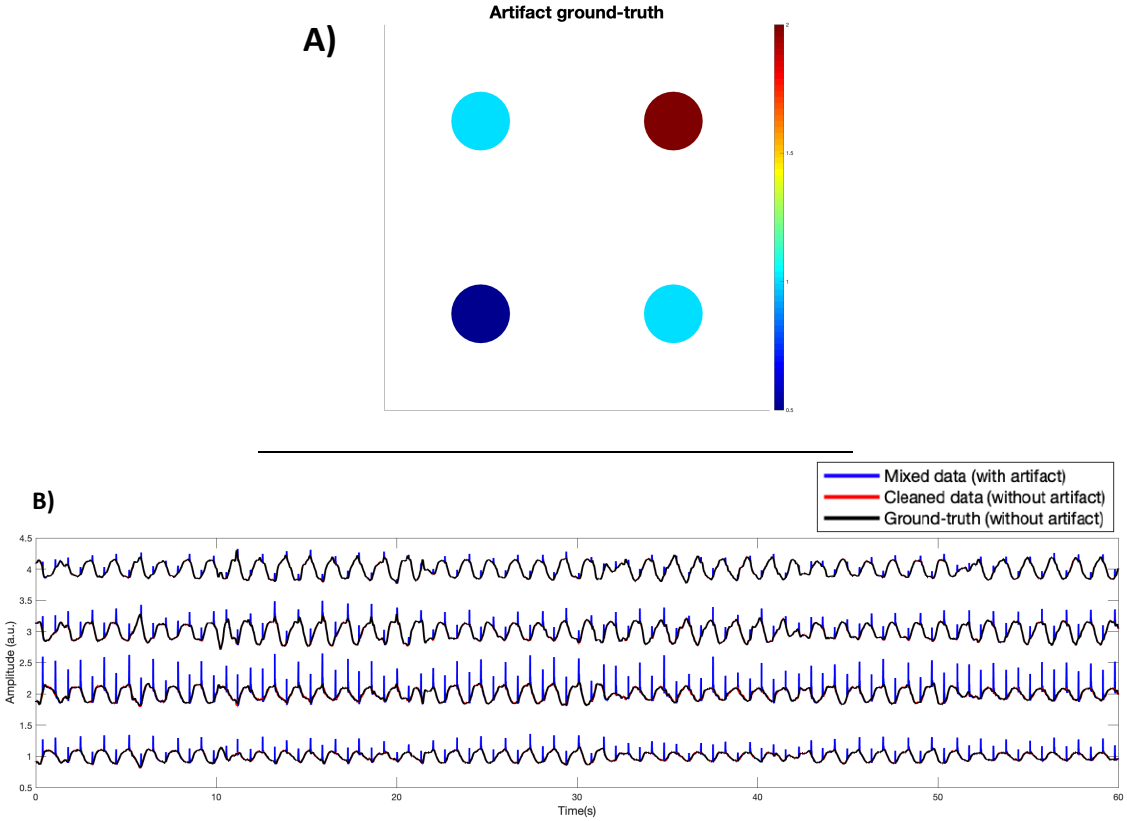


Figure 8. 3. Assessment of the ICA performance for artifact removal. A) *Forward model.* The topography of the forward model reflects maximum projection of the tapping data to the front-right sensor, and minimum projection to the back-left sensor. This spatial pattern was empirically observed in real data. B) *Artifact removal.* The mixed data (blue line, tapping timeseries added to the sensors activation) reflect the spatial distribution of the forward model. When comparing the clean data (red line, after ICA removal) with the ground-truth (black line, mock data before adding the artifact contribution), it is evident that the two lines overlap. This shows that the approach is effective in automatically identifying and removing the target artifactual component from the force plate sensor data.

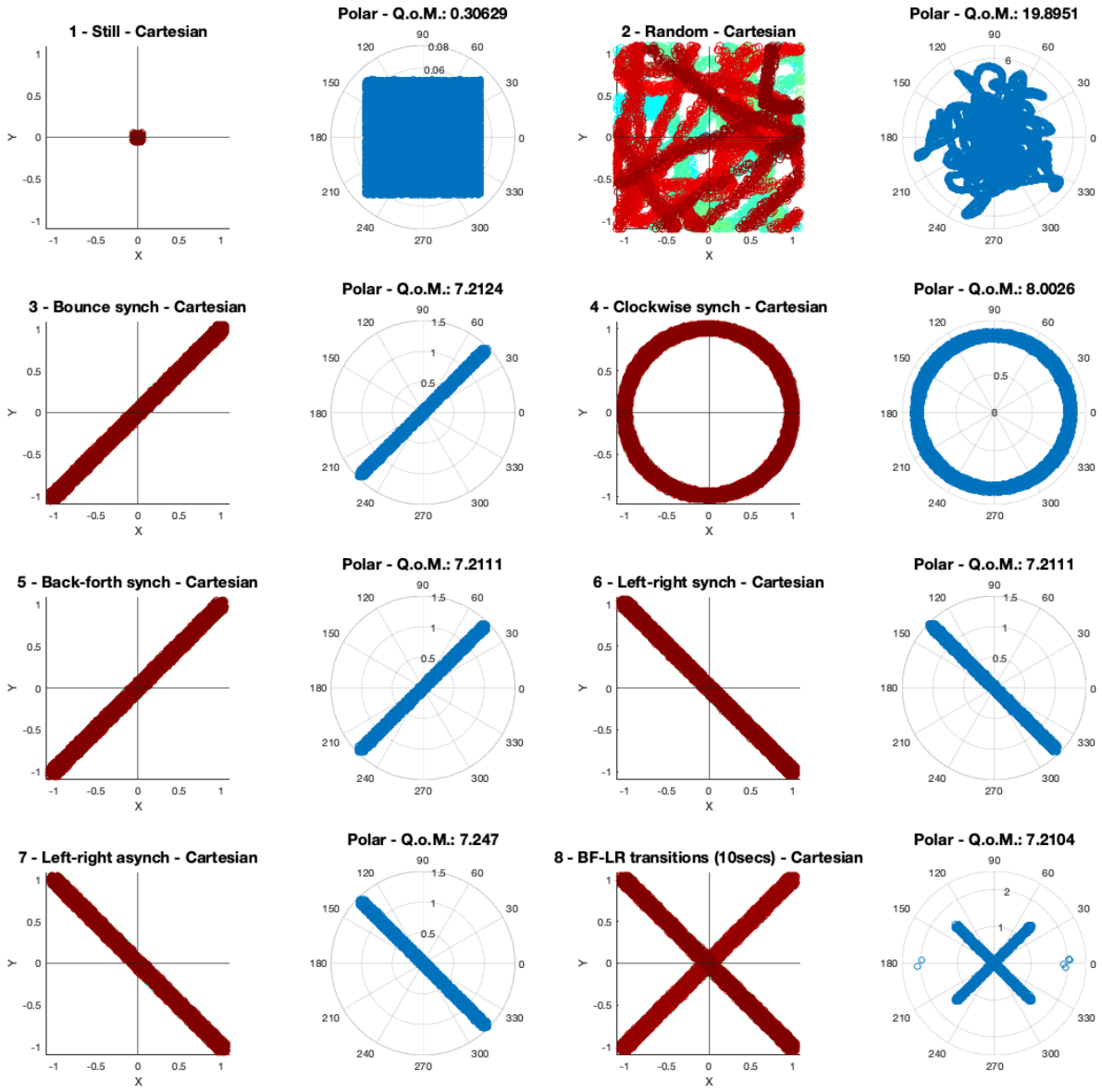


Figure 8. 4. Simulated data in Cartesian and polar coordinates. For each behavior, data points are represented in Cartesian (to the left) and polar (to the right) coordinates. In Cartesian coordinates, the time dimension is color-mapped from blue to red, to verify that the spatial distribution reflects the sensors' activation over time (see *Figure 8.1*). QoM is reported for each behavior, as computed from polar coordinates.

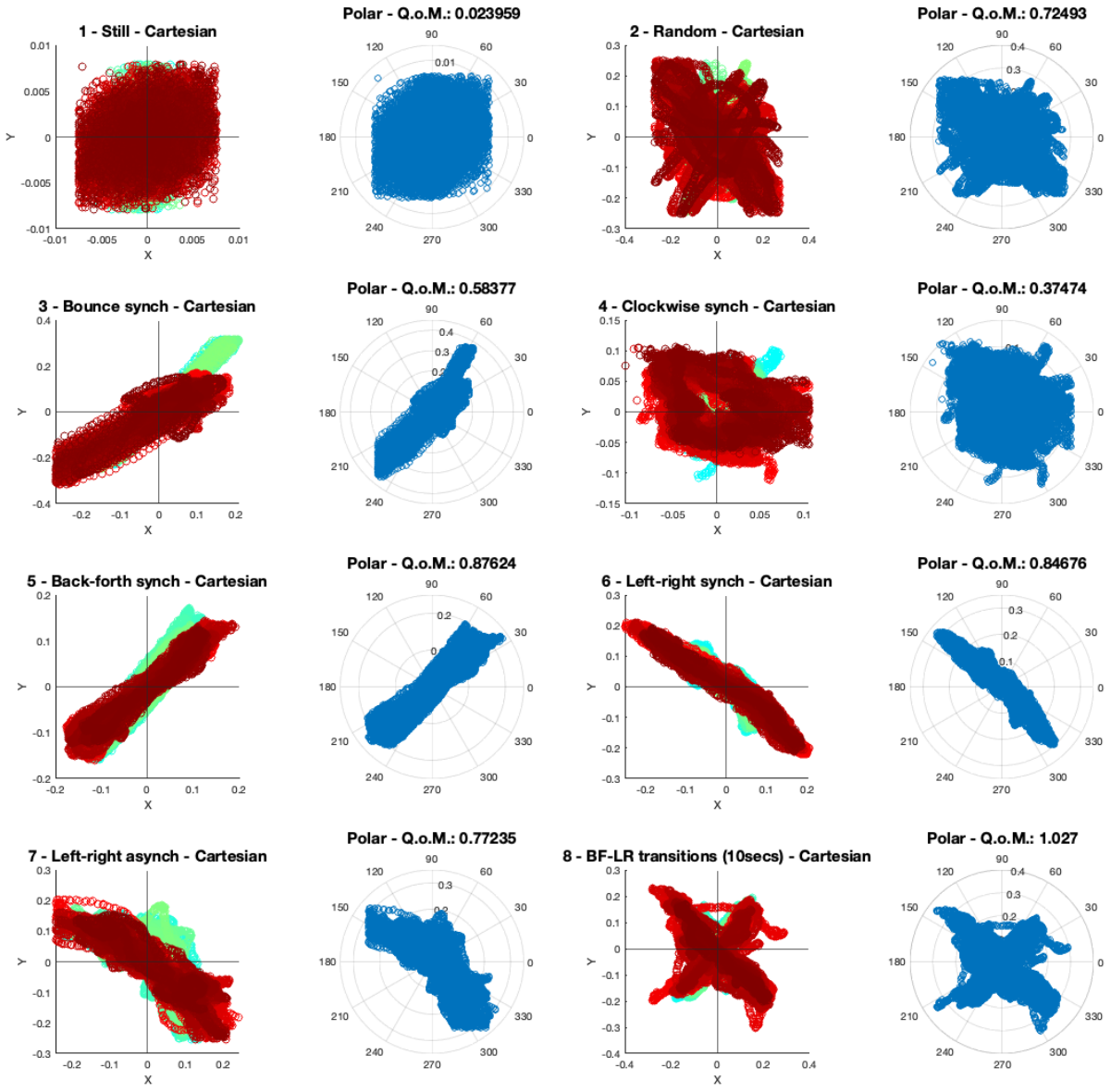


Figure 8. 5. Mock data in Cartesian and polar coordinates. For each behavior, data points are represented in Cartesian (to the left) and polar (to the right) coordinates. In Cartesian coordinates, the time dimension is color-mapped from blue to red, to verify that the spatial distribution reflects the sensors' activation over time (see *Figure 8.2*). QoM is reported for each behavior, as computed from polar coordinates.

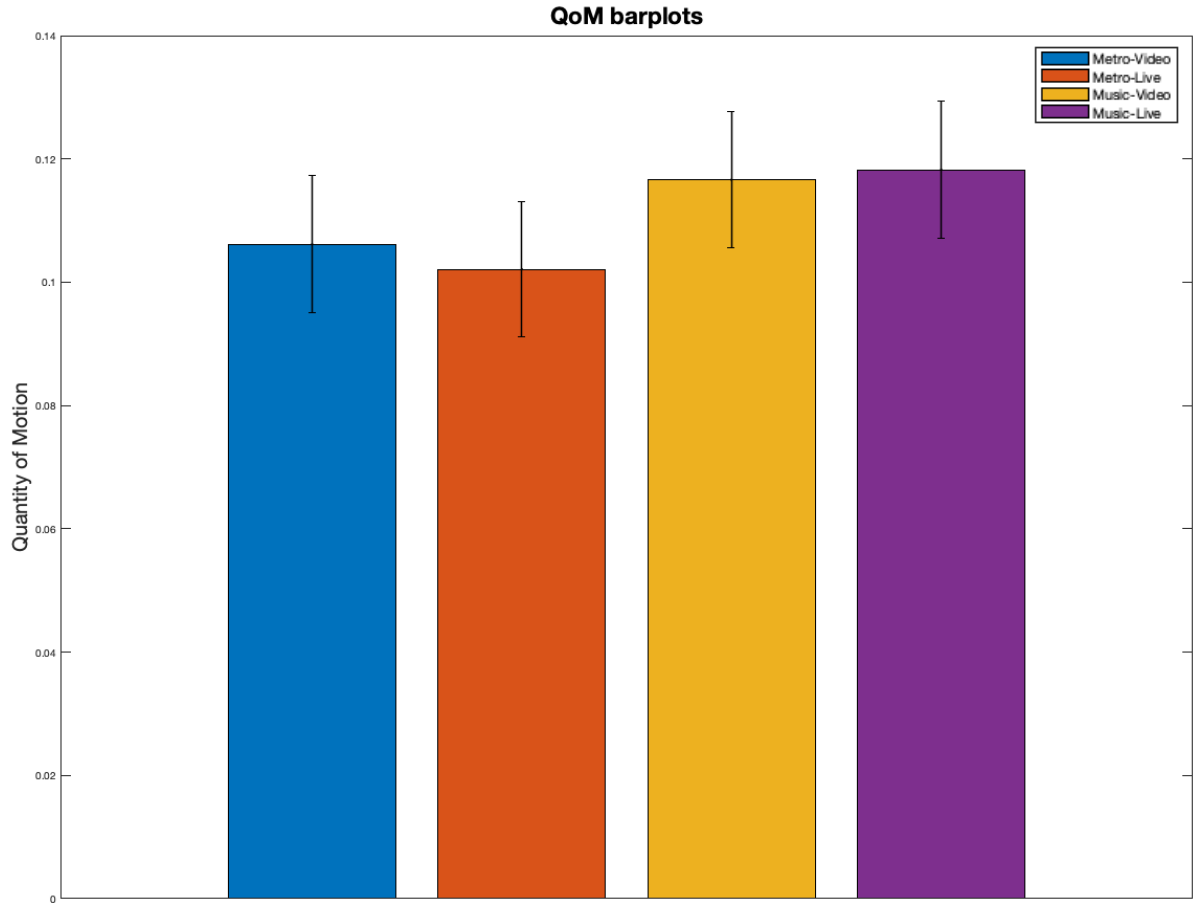


Figure 8. 6. Average quantity of motion (QoM) and standard errors of the mean (SEM). A first visual inspection suggests that Music conditions resulted in greater QoM, and that Metronome condition were more susceptible to the physical presence of the musician. In other words, participants exhibited more body sway when synchronizing with music as compared to metronomes, regardless of the physical presence of the musician. When synchronizing with metronomes, the video performance of the musician may have induced more movement as compared to the live performance. We address a 2 x 2 ANOVA as first statistical test to assess the significance of main and interaction effects of Stimulus and Social factors, followed by the inclusion of MMSE, age, and other demographical variables as additional predictors.

2. Recurrence quantification analysis (RQA)

Dimensionality reduction via PCA

In reference to the introduction of Independent Component Analysis (ICA), having a multivariate dataset allows for the application of multivariate methodologies. Our objective at this stage was to extract meaningful spatiotemporal behavioral patterns from the sensors data, while concurrently reducing the dataset's dimensionality. Principal Component Analysis (PCA) is an appropriate strategy for our goal. The idea was to reconstruct behavior as a source of data variance, based on covariance patterns in sensor activation over time. The anticipated result of the analysis is to derive one or two single timeseries that could account for the majority of the data variance, which could then be used as input for subsequent stages of the analysis pipeline (specifically, RQA and JRQA).

We implemented PCA on simulated and mock datasets of stereotyped behaviors with the following objectives:

1. Assess whether a single behavioral pattern can be adequately represented by a single component.

2. Evaluate the invariance of source separation to temporal consistency, spatial patterns, and transitions within the behavior.

3. Explore dependencies on specific features inherent to the device and actual human behavior.

We initiated by centering the input sensor data. Subsequently, a channel-by-channel covariance matrix (*C*) was computed. The eigenvectors (*W*) and eigenvalues (Λ) were determined via eigendecomposition of the covariance matrix. The following eigenvalue equation was solved for the weights:

$$W\Lambda = CW$$

Subsequent to this, the eigenvectors (W) were sorted in descending order based on the magnitude of their corresponding eigenvalues (Λ). The data matrix (X) was then multiplied by the first of the sorted eigenvectors to derive the principal components (PCs) scores, also known as activation timeseries (*Y*).

$$Y = w^T X$$

Additionally, the activation maps or spatial patterns of the PCs were computed as the inverse of the eigenvectors matrix (W^{-1}).

Ultimately, to eliminate the dependence on the data scale, we normalized the eigenvalues into percentage units representing explained variance. This normalization enables the evaluation and comparison of the PCA quality across different datasets and experimental conditions. Besides visually inspecting the PC scores over time, the formal evaluation of PCA is based on the eigenspectrum, which is the distribution of variance explained by each PC. As mentioned earlier, our goal was for the first component's eigenvalue to explain the majority of the data matrix's variance. This would allow us to confidently compress the data to the single associated PC for subsequent analysis. *Figure 8.5* illustrates the data in sensor and PC space, along with the PC distributions and eigenspectrum, for both simulated and mock data.

Assessing PCA on stereotyped behaviors recorded by a four-sensor force plate yielded the following insights:

1. When sensor data comprise a single behavioral pattern (behaviors #3, #5, #6, #7), this is suitably explained by a single PC in both simulated and mock data. Exceptions include circular motion (behavior #4) and transitions across orthogonal spatial axes (behavior #8), where two components prominently appear in the eigenspectrum. In these cases, a state-space defined by these orthogonal dimensions would offer a more accurate representation as compared to PC activation over time. Behaviors #1 and #2 (white and pink noise, respectively) exhibited uniform or near-uniform eigenspectra, with the latter revealing some decay due to its autocorrelation structure.

2. If a single pattern is consistently dominant in the data, source separation is invariant to both spatial distribution and temporal consistency. Remarkably, it makes no difference whether periodic pressure affects all sensors simultaneously (behavior #3) or pairs of sensors in back-and-forth (behavior #5) or left-to-right motion (behavior #6). Even the temporal stability of the pattern does not influence the result, as long as the covariance across sensors remains consistent over time (behavior #7). Systematic transitions from one spatial pattern to another, however,

lead to the extraction of two components with equivalent explanatory power, whose relative activations are timed with the transitions (behavior #8).

3. The comparison of simulated and mock data reveals a remarkable similarity in the effectiveness of PCA when motion follows a distinct pattern (behaviors #3 to #8). However, it becomes clear that noise at rest is to some extent correlated across sensors, resulting in a less uniform eigenspectrum for mock data compared to simulated data (behavior #1). It is plausible that the experimenter's micro-movements contributed to this while attempting to sit still, introducing further correlation across sensors. Notably, the comparison produced a steeper decay of the eigenspectrum in the case of random motion (behavior #2). Real human motion not only implies an autocorrelation structure (at each time point, the value of each sensor is dependent on previous data points), but the physical constraints due to the mechanical coupling of the experimental setup and the human body impose a correlational structure across sensors, which is reflected in the covariance matrix.

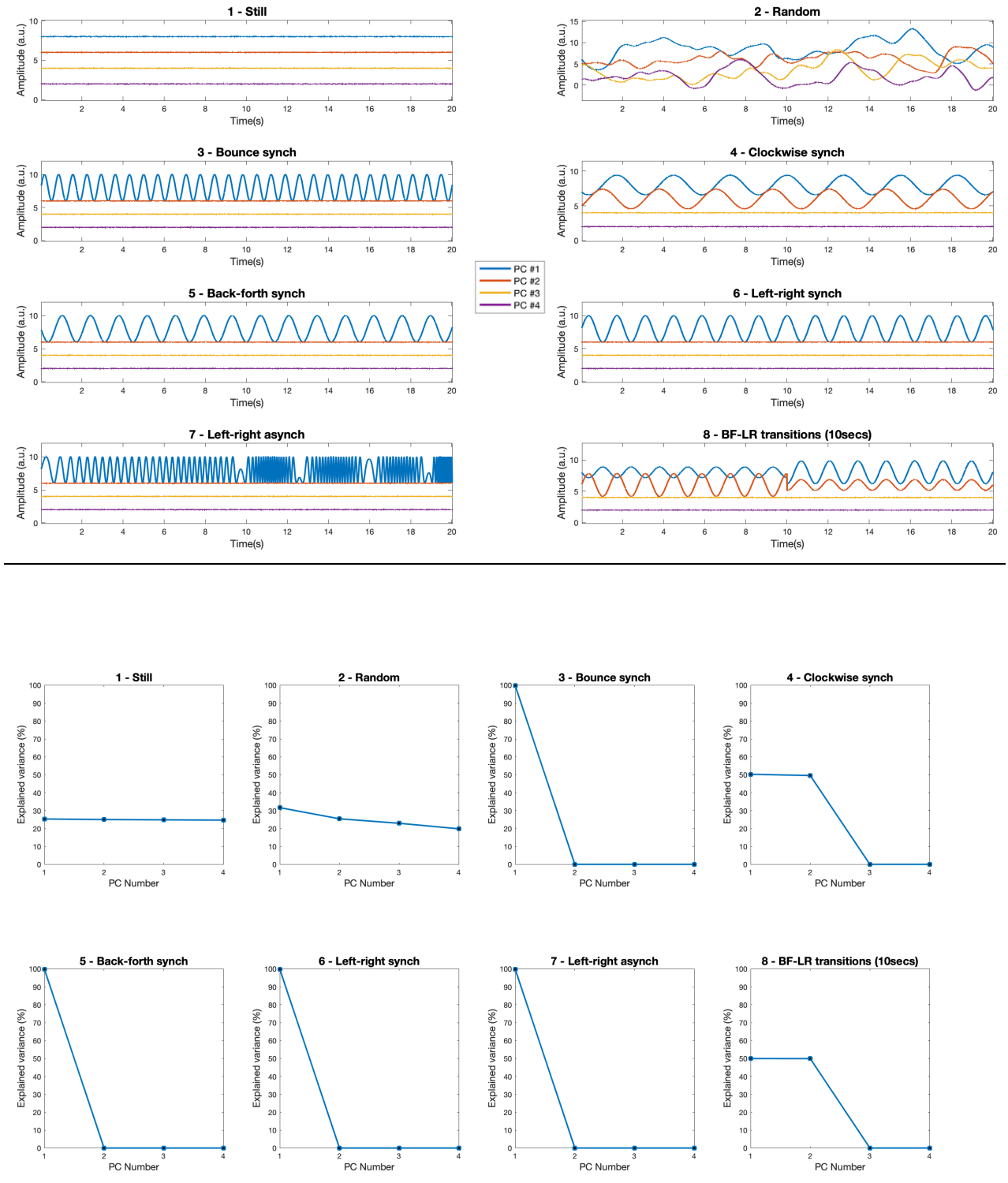


Figure 8. 7. Principal components (PCs) scores and eigenspectra of simulated data.

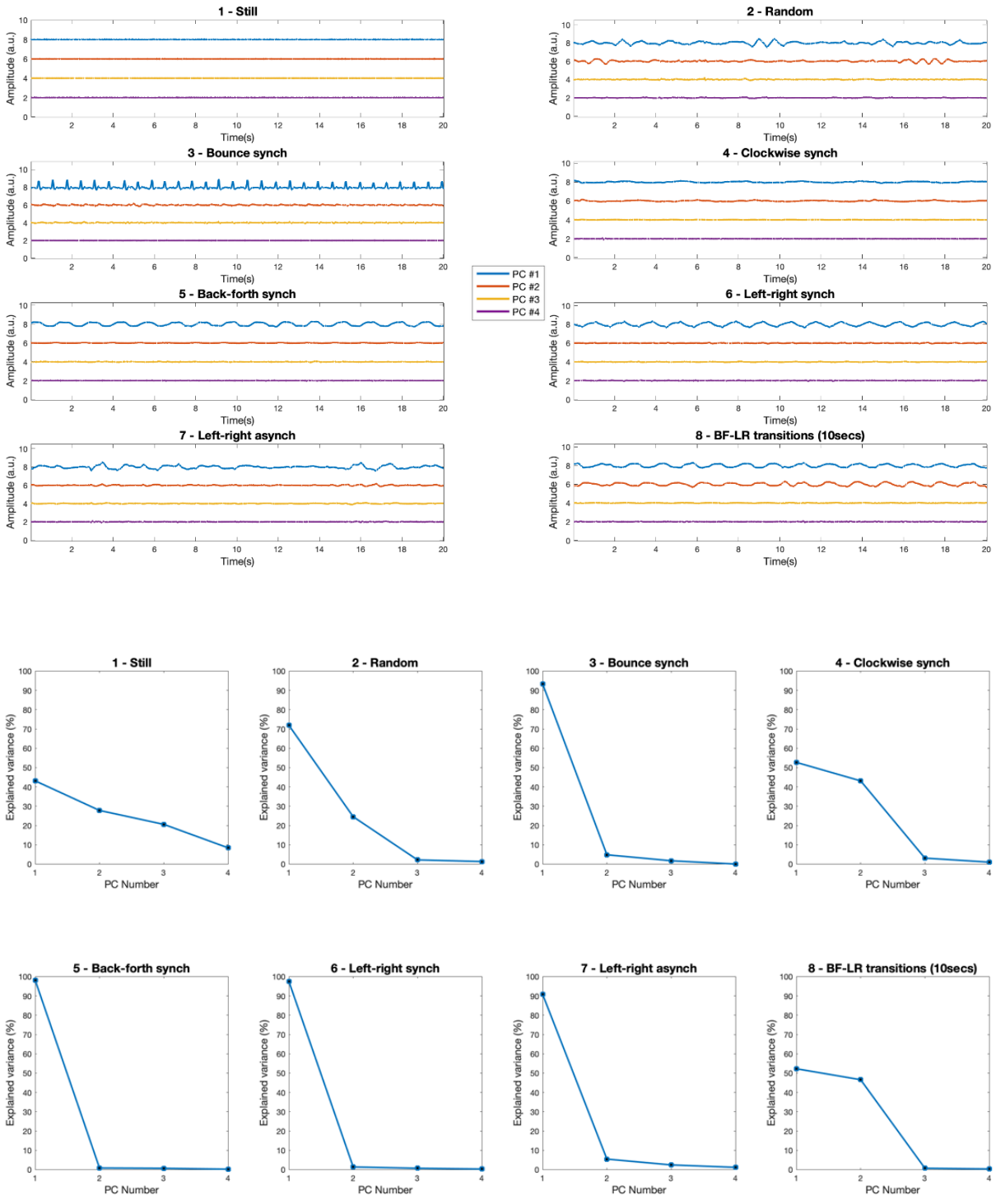


Figure 8. 8. Principal components (PCs) scores and eigenspectra of mock data.

Complexity measures from recurrence plots (RPs)

Once the data are reduced to a single timeseries, it becomes feasible to conduct timeseries analysis. In this study, we elected to use Recurrence Quantification Analysis (RQA) to investigate the recurrences in patients' body sway during musical interaction (Demos & Chaffin, 2017; Demos, Chaffin, & Kant, 2014; Demos et al., 2018). As discussed in Chapters 2-4, RQA is based on the embedding of timeseries in phase space and subsequent computation of Recurrence Plots (RPs). These square matrices quantify the recurrence of states in a dynamical system, based on its trajectory in phase space (Eckmann et al., 1995). For the mathematical formulation of time-delayed embedding (Takens, 1981) and the computation of RPs, we refer back to those Chapters. In this section, we will present the two complexity measures computed via RQA that we deemed most relevant to this study. These measures allow to move beyond the visual interpretations provided by RPs by quantifying the small-scale structures embedded in the timeseries.

These measures constitute RQA in the strict sense (Marwan et al., 2007), and are derived from the density of recurrence points and the diagonal and vertical line structures within the RP (Marwan, Wessel, Meyerfeldt, Schirdewan, & Kurths, 2002; Webber & Zbilut, 1994; Zbilut & Webber, 1992). The vertical structures in the RP are connected with intermittent and laminar states of a system, while the diagonal lines relate to its predictability. The most straightforward measure within RQA is the recurrence rate (RR), also known as percent recurrences:

$$RR(\varepsilon) = \frac{1}{N^2} \sum_{i,j=1}^{N} R_{i,j}(\varepsilon)$$

RR quantifies the density of recurrence points within the RP, consisting of the proportion of time a system spends either locked in or revisiting a particular state within the boundaries of a threshold ε, defined as the radius of the phase space. Furthermore, it offers an estimation of the likelihood that a system's state will recur in the future.

The second measure in this analysis is the determinism of the system (DET), and is derived from the histogram $P(\varepsilon, l)$ of diagonal lines of a given length l.

$$DET = \frac{\sum_{l=l_{min}}^{N} lP(l)}{\sum_{l=1}^{N} lP(l)}$$

Uncorrelated or weakly correlated processes, whether stochastic or chaotic in nature, result in few or very short diagonals, while deterministic processes result in longer diagonals and fewer isolated or vertically aligned recurrence points (Marwan et al., 2007). Therefore, the ratio of recurrence points forming diagonal structures (of at least length l_{min}) to all recurrence points provides a measure of determinism in the system. The threshold l_{min} excludes the diagonal lines formed by the tangential motion of the phase space trajectory. It is important to note that both RR and DET are proportions, and they are normalized on a scale ranging from 0 to 1.

These measures offer insights that align with the linear and circular methods conventionally employed for SMS quantification. For instance, a high RR may suggest high autocorrelation in the timeseries, and a high DET relates to a high vector length, since consistent synchronization with an isochronous stimulus results in deterministic behavior. However, while these measures may intersect in certain aspects, RQA provides additional insights into system dynamics that linear methods may overlook, especially in instances where assumptions of stationarity and linear relationships are violated. The approach is explicitly designed to capture nonlinear dynamics and intricate data patterns.

We hereby present RR and DET calculations based on 164 participants (N = 164), because 30 participants had missing conditions or exhibited major issues in the recording which made the analyses unfeasible. The activation timeseries of the first principal component (PC) was used as input for the analysis, after down sampling by a factor of 10 for computational feasibility. The parameters for time-delayed embedding: delay τ = 12.635, dimensions m = 5.141, and radius ε = 10% of the phase space diameter, were adjusted according to the same criteria outlined in Chapters 2-4. The only difference is that for this dataset, parameters were individually optimized for each participant by averaging values resulting from the four experimental conditions. This decision was informed by the observation that using the same average parameters on different patients led to extreme floor and ceiling effects for both measures. This was mostly due to the different sensitivities to the threshold of the radius ε in the phase space.

It is plausible that this sensitivity difference is due to the high variability in the population of interest. Factors such as responsiveness to auditory stimuli, interaction with the musician, and mobility could influence this, especially considering that full body movement was not instructed during the SMS task. However, we also observed that

recording clusters conducted in different timeframes exhibited distinct SNRs. The factors underlying these differences remain unknown to the data analyst, but are likely to influence the parameter tuning. Adjusting the parameters for each participant is not inherently problematic, as long as they are kept consistent across experimental conditions. However, this necessitates treating participants as random effects in the statistical models, given that the intercept will vary in line with the parameters. *Figure 8.9* shows the average values of RR and DET in the different experimental conditions. Mean and standard deviation are reported in *Table 8.1*.

The bar plots suggest that participants exhibited higher DET in Music conditions, as compared to Metronome conditions. This indicates that spontaneous body sway becomes more predictable when the SMS task is paced by a musical piece rather than an isochronous metronome. This is particularly interesting, given that metronomes have repeatedly shown to better facilitate consistency in SMS in this population. Our inspection suggests a dissociation between the mechanisms underlying intentional finger-tapping and spontaneous body sway with respect to the rhythmic stimulus, which points to the involvement of distinct but possibly overlapping systems. Differences in RR across conditions are less evident upon visual inspection.

Like in the case of QoM, we address a 2 x 2 ANOVA as first statistical test to assess the significance of main and interaction effects of Stimulus and Social factors, for both RQA measures. This first tests should be followed by the inclusion of MMSE, age, and other demographical variables as additional predictors.

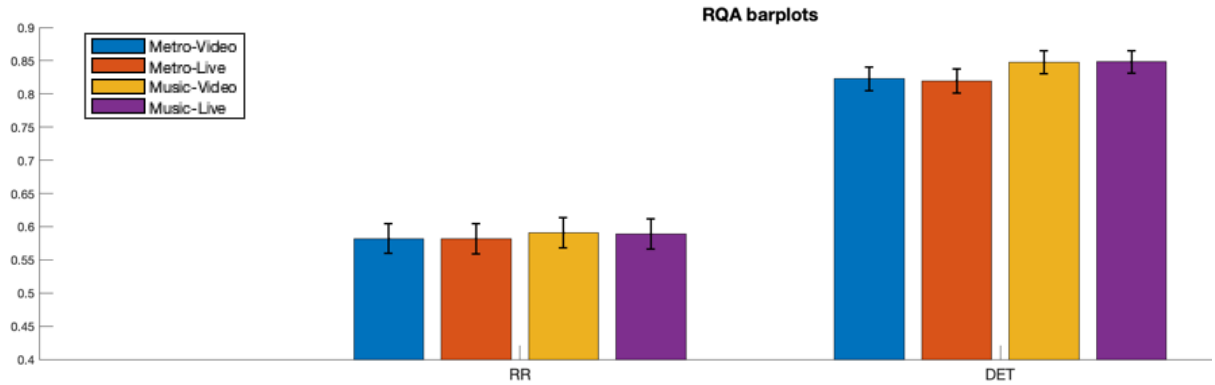


Figure 8. 9. Average recurrence quantification analysis (RQA) measures and SEM.

3. Joint Recurrence quantification analysis (JRQA)

RQA can be expanded to a dyadic setting, encompassing the mirroring between participant's and musician's body movements during the interaction, resulting in the co-occurrence of spatiotemporal patterns. JRQA consists of computing the same RQA measures from Joint Recurrence Plots (JRPs), although it should be noted that some interpretations might not translate directly in the dyadic version.

The extension of RQA to multiple musical performances can be declined into two approaches: Cross Recurrence Quantification Analysis (CRQA) and Joint Recurrence Quantification Analysis (JRQA), each answering a distinct query (Demos & Chaffin, 2017). CRQA seeks to identify whether a specific pattern is repeated across two performances, disregarding the location within the piece. Conversely, JRQA investigates if certain patterns are replicated at the same time in both individual behaviors, thus focusing on the co-occurrence of patterns in time. Since it is based on the co-occurrence, while disregarding the particular pattern taking place, JRQA is in principle blind to the layout of the force plate sensors. In practice, JRQA overlays two (or more) RQA plots, pinpointing when recurrence occur in the same entry of both RPs. In other words, it is computed as the Hadamard product between two RPs. The overlaying recurrent points might reflect distinct movements, but they indicate that the performer executed a recurrent action at the same point in each performance. This results in a new plot that shows where both RPs exhibit recurrent patterns. Therefore, JRQA offers a solution to identifying similarities in recurrence across performances, rather than similarities in the movements themselves. We argue that this is a more informative approach, because it is sensitive to coordination in time while being tolerant to spatial differences between individual movements. We can expect these to be considerable, given that we are comparing an elderly NCD population to a young healthy musician.

It should be kept in mind that both CRQA and JRQA require the performances being analyzed together to operate on the same time scale. In addition to starting with timeseries of equal size, the embedding parameters τ and m must be consistently employed to ensure that the resulting phase spaces and RPs are also of the same size. For each participant in our study, the activation timeseries of the first principal component (PC), extracted from the force plate data, was used as the first input. The same parameters as the individual RQA were applied. The corresponding PC of the musician served as the second input. The embedding parameters applied to the participant were also applied to the musician to maintain consistency. It is important to note that force plate data for the musician were only available in Live conditions. Therefore, conducting JRQA in Video conditions was not feasible in this particular study.

Even for JRQA, results suggest that the co-occurrence of recurrences produced higher DET in the Music Live conditions as compared to Metronome Live. This indicates that the dyadic system formed by the participant and the musician becomes more predictable when their live interaction is mediated by a musical piece. We address once again the same statistical tests pointed out for previous measures.

RQA and JRQA measures were computed with the *CRP Toolbox* for Matlab (Marwan & Kurths, 2002). The descriptive statistics used to generate the bar plots for this dataset are reported in *Table 8.1*.

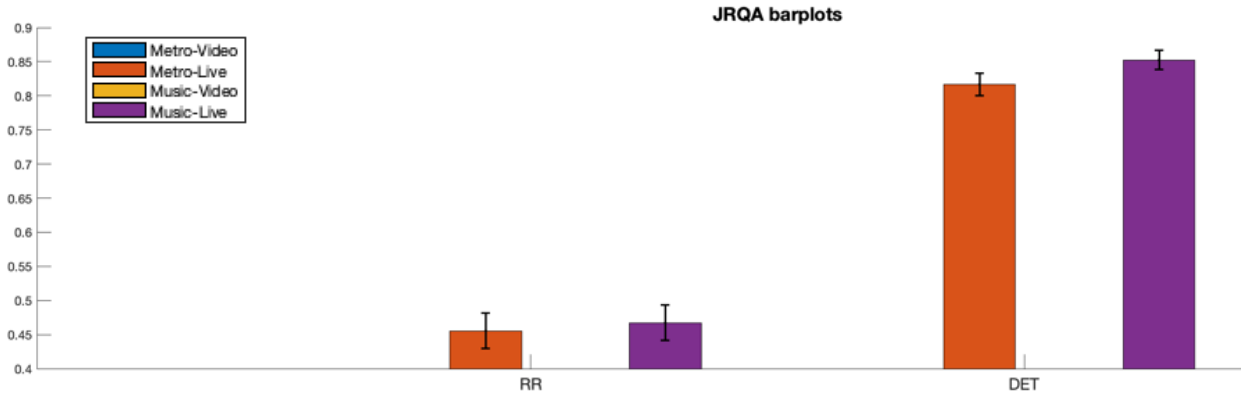


Figure 8. 10. Average joint recurrence quantification analysis (JRQA) measures and SEM.

		Metronomes		Music	
		Video	Live	Video	Live
Mean	QoM	0,1062	0,1021	0,1167	0,1183
	RQA_RR	0,5818	0,5818	0,5906	0,589
	RQA_DET	0,8227	0,8198	0,8476	0,8482
	JRQA_RR	N.A.	0,4555	N.A.	0,4674
	JRQA_DET	N.A.	0,8168	N.A.	0,8526
Standard Deviation	QoM	0,1420	0,1406	0,1408	0,143
	RQA_RR	0,2863	0,2898	0,2907	0,2902
	RQA_DET	0,2271	0,2323	0,2176	0,2131
	JRQA_RR	N.A.	0,3317	N.A.	0,3325
	JRQA_DET	N.A.	0,2068	N.A.	0,184

Table 8. 1. Descriptive statistics for QoM, RQA, and JRQA.

Dataset #2 – Dynamic adaptation to tempo-changes (von Schnehen et al., *in preparation*)

An ongoing study by von Scheenen et al. (*in preparation*) aims to investigate the effects of NCDs on sensory motor synchronization, specifically the ability to coordinate motor rhythms with external rhythms at different levels of complexity in their temporal structure. Inducing synchronization errors with perturbations aimed at eliciting error-correction responses provides a great opportunity to zoom-in into the dynamics of patients' adaptive responses to the new tempo within critical time windows. As highlighted in the major methodological principles of Chapter 1, such an approach is key to investigate the dynamics and the control mechanisms underlying SMS. Because error correction is a crucial component of minimizing asynchrony with a rhythmic stimulus, systematically perturbing isochrony is a valuable approach to induce errors and study corrective responses (Jantzen et al., 2018; Praamstra et al., 2003; B. H. Repp, 2001a, 2001b; Bruno H. Repp, 2000; Bruno H. Repp & Keller, 2004), providing insight into the underlying timing mechanisms. In turn, this knowledge is necessary to understand the potential mechanisms underlying the effectiveness of music-based interventions on NCD (von Schnehen et al., 2022). Based on neuropsychological assessment, the sample of participants was divided into No NCD, Minor NCD, and Major NCD groups, which allowed for the assessment of the degeneration of dynamics throughout the stages of the disease.

The perturbation-based approach in this study is not dissimilar to the tempo-change perturbations implemented with young, healthy participants in Chapter 7, albeit adapted to the population of interest. However, we argue that to maximize the inferences drawn from the data, it is essential to utilize analysis methods that align with the methodological rationale of eliciting dynamics. Specifically, it is crucial to move beyond global measures of central tendency and dispersion of SMS, which form the basis of descriptive circular statistics (Berens, 2009; Fisher, 1995). Given the temporal structure of the task, the dataset is suitable for processing as a behavioral Event-Related Frequency Adjustment (ERFA) to model the frequency change in finger-tapping behavior in response to perturbations (see Chapter 7). Instantaneous frequency timeseries were estimated from the onsets of the finger-taps and modeled with a set of non-linear functions that best capture the behavior observed in empirical curves. The next section provides a synthesis of the methods currently adopted in the ongoing study, followed by two data analysis approaches developed in our new analysis framework, along with their respective preliminary results.

Study design

Participants

The study recruited a sample of 51 older patients (N = 51; mean age = 81.15 years, std = 6.97) who visited a geriatric hospital for a multi-disciplinary evaluation. Following the evaluation, the patients were given a diagnosis, which the experimenter was blind to at the time of testing. The participants were further divided into two groups based on their Mini-Mental State Examination (MMSE) score, with scores of 25 and above indicating healthy aging and scores below 25 suggesting early dementia.

Stimuli and Task

The study aimed to investigate the impact of neurocognitive disorders on SMS with stable and perturbated rhythmic sequences. The experiment employed a paradigm that involved participants tapping to the rhythm of a musical piece or a metronome. Each participant underwent eight trials with different conditions: four types of temporal structure (constant tempo at 89 BPM, constant tempo at 81 BPM, shifting between the two tempi starting from 89BPM, shifting between the two tempi starting from 81 BPM), and two types of auditory context (musical piece or metronome). The trials consisted of 75-second-long stimuli, with sudden accelerations or decelerations every 15 seconds for the shifting tempo conditions. The task was designed to accommodate individuals with moderate to severe dementia by having them seated across a life-sized screen, displaying another person performing the task with them. This approach aimed to motivate the participants and encourage them to engage with the task consistently.

Experimental design

The study aims at examining several within-subject dependent variables, including the type of auditory context (music or metronome), tempo (89 or 81 BPM), presence or absence of tempo changes, and the effect of interval (performance differences at the beginning of the trial versus later). Between-subject variables included diagnosis (No NCD, Minor NCD, Major NCD) and a continuous MMSE score. The dependent variable was tapping consistency, defined as the variability of the timing difference between taps and pacing stimuli or taps and beats. Consistency was expressed as a number between one and zero, with one indicating optimal performance and zero representing random tapping with no relation to the stimulus.

Data analysis

The subsequent analyses were performed on the onset data extracted from the finger-tapping timeseries recorded during the experimental sessions. The first step involves the computation of instantaneous frequency. Following this, the analyses diverged into two distinct approaches.

Processing: Instantaneous frequency computation

The intervals between finger-tapping onsets were linearly interpolated from 0 to 2π at a sampling rate of 1 kHz, yielding an estimate of the finger-taps phase with a millisecond temporal resolution. Instantaneous frequency timeseries were then computed as the first derivative of the unwrapped phase angles timeseries, and converted to Hz in order to express frequency over time:

$$Hz_t = \frac{s(\phi_t - \phi_{t-1})}{2\pi}$$

where *s* indicates the data sampling rate, and ϕ_t indicates the (unwrapped) phase angle at time *t*.

The presence of missed taps in the timeseries poses a significant challenge, particularly with this population when the task is executed over an extended period. While the use of circular statistics has been suggested to be robust to this issue (Ghilain et al., 2020), this claim is debatable. Similar to linear measures of asynchrony, circular measures based on relative phase are calculated as the difference between the taps and the reference stimulus sequence. Rescaling and normalizing to radians do not alter the fact that some onsets may be absent in the dataset. Therefore, this problem must be addressed during the data processing phase. As in all previous chapters, we tackled this issue by employing linear interpolation of the onsets timeseries at a defined sampling rate, which ensures alignment across both behavioral and stimulation timeseries. With our approach, we should bear in mind that eventual missed taps would result in a drop of instantaneous frequency. At the group level, these are presumed to occur at random times during the task and would therefore cancel out when performing averages and curve fitting.

Related to this, a critical point of divergence between this approach and others previously proposed pertains to the handling and definition of outliers. The calculation of relative phase between finger-taps and beats implies a sort of pre-selection of the best pairs

because, in instances of omissions or double-taps within beats, the nearest neighbor to the 'correct' tap is chosen. While this can be viewed as a form of data cleaning, it comes at the cost of losing information about the actual behavior, introducing a bias that neglects dynamics causing abnormalities. In contrast, instantaneous frequency is reference-free in the sense that it isn't defined based on the stimulation timeseries. A participant who misses a tap or unexpectedly taps twice in-between beats will produce a dynamic in the timeseries. In the not unlikely scenario of a participant consistently tapping to harmonics or sub-harmonics of the stimulation frequency, an offset will be produced in the timeseries. Although these situations may be challenging to manage, they provide valuable information that should not be ignored at this stage of processing. This information can be compensated for later through normalization procedures or alternative procedures for outlier removal. As demonstrated in the following paragraphs, in absence of a singular criterion for identifying outliers based on instantaneous frequency alone, all timeseries were modeled using the approaches presented below. Outliers will be defined later on, based on the distribution of the models' parameters.

Approach #1: Global frequency modulation

The perturbation task used in this study was tailored for the NCD population. Consequently, unlike the ERFA study discussed in Chapter 7, we could not rely on a high number of critical events to optimize the signal-to-noise ratio (SNR) by averaging numerous responses to individual events. However, tempo changes occurred here in evenly spaced, lengthy steps. This temporal structure resulted in an expected periodic modulation of the instantaneous frequency timeseries over 15-second periods. *Figure 8.11* illustrates how the empirical timeseries effectively tracked the expected pattern. Specifically, in 'Fast' and 'Slow' steady tempo conditions, the instantaneous frequency fluctuates around the stimulation frequency. In contrast, in 'W' and 'M' tempo-change conditions, it traces the changes in mirrored patterns depending on the initial tempo. Given the consistency of this pattern over the macro-temporal structure, we propose an analysis approach applied to the longitudinal format of the data. This approach compensates for the limited number of critical events of interest occurring during the trial. A visual inspection of the data suggested that fitting a sine wave to the instantaneous frequency curves could be an effective method to model the data, given the oscillating changes in frequency following periodic tempo changes.

The sine wave function was defined by the following equation, previously introduced in Chapter 1:

$$y(t) = A * sin(2\pi * f * t + \varphi) + C$$

This yielded four parameters for estimation: amplitude (A), frequency (ω), phase (ϕ), and offset (C). Utilizing the *sineFit()* function in Matlab (Seibold, 2021), we fitted the curves through an optimization process that aimed to minimize the sum of squared errors between the fitted sine wave and the data. The initial 15 seconds were omitted from the fit, as no critical event was occurring during this period, and it gave participants some time to synchronize their behavior. Fitted curves are depicted in *Figure 8.12,* and grand-averages are shown in *Figure 8.13*. Vertical lines represent the onset of tempo changes, while horizontal dashed lines represent the expected 10% range spanned by tempo-changes.

As seen in tempo-change conditions, the best fitting sine wave consistently exhibits periodic fluctuations that track the expected pattern of perturbations. Conversely, in steady conditions, where random fluctuations due to tapping variability are not expected to be periodic, the best fitting sine wave has a consistent offset at the average expected tempo. In these instances, in absence of major artifacts, inherent noise in the movement should result in random frequency and phase, along with a small amplitude. Given the context of the task, we propose the following interpretations for the parameters that define these curves:

- Amplitude (A). This represents the magnitude of adaptation to the periodic tempo change and indicates the size of the effect. We expected a 10% change with respect to the steady tempo.
- Frequency (f): This represents the rate of tempo change and indicates the presence of the effect. Given one tempo change every 15 seconds, we expected a frequency of 0.0333 Hz.
- Phase (ϕ): This represents an offset along the time axis and indicates overall reactivity to perturbations. We interpret a phase of 0 rad as perfect synchrony with the tempo change.
- Offset (C): This represents the average value of the sine wave and indicates the average tempo maintained throughout the task. The expected offset would be 1.483 Hz for Fast steady tempo, 1.350 Hz for Slow steady tempo, and 1.417 Hz for both tempo-change conditions (as the average of Fast and Slow).

Importantly, while the information gathered during the entire duration of the task is condensed into individual values, these values remain sensitive to specific aspects related to the adaptive behavioral dynamics of the task's temporal structure. We can illustrate how the sinusoidal fit compares to circular statistics and associated angular measures by considering some hypothetical scenarios. For instance, because adaptation takes some time to occur, the perturbation would invariably reduce synchronization during critical time windows, which would then affect the global measure. Observing lower synchronization consistency in tempo-changing conditions wouldn't necessarily indicate that the participant wasn't adapting to the change, only that synchronization was temporarily destabilized. The same holds true for the mean relative-phase angle, given that a systematic lag is anticipated in response to periodic perturbations. Consequently, lower accuracy and consistency wouldn't necessarily indicate a lack of adaptation.

On the other hand, the four parameters of the sinusoidal fit provide insight into specific temporal dimensions that directly correspond to the pattern of tempo change. The distributions of these parameters across experimental conditions are depicted in *Figure 8.14*. We identified 10 participants deviating by more than 3 standard deviations from the mean values of either offset or amplitude parameters. These were considered outliers and removed.

As anticipated, in steady conditions, frequency and phase are uniformly distributed because no periodic changes are involved, amplitude is consistently low, and the mean offset corresponds to the stimulation tempo. This indicates that participants were overall synchronizing to the steady stimulus. In the tempo-change conditions, frequency is densely clustered around the rate of tempo change, phase offset is concentrated around a specific lag (depending on the 'M' or 'W' pattern), amplitude is consistently high, and the offset is situated around the average of the two alternating tempos.

For subsequent statistical analyses, the parameter estimates from steady tempo conditions will be used as baselines to evaluate the systematicity of variations resulting from the temporal pattern of tempo changes. These calculated measures enable statistical comparisons across metronome and music conditions, as well as across participant groups. These statistical analyses lie beyond the scope of this Chapter. The descriptive statistics of the sine parameters' distributions are reported in *Table 8.2*.

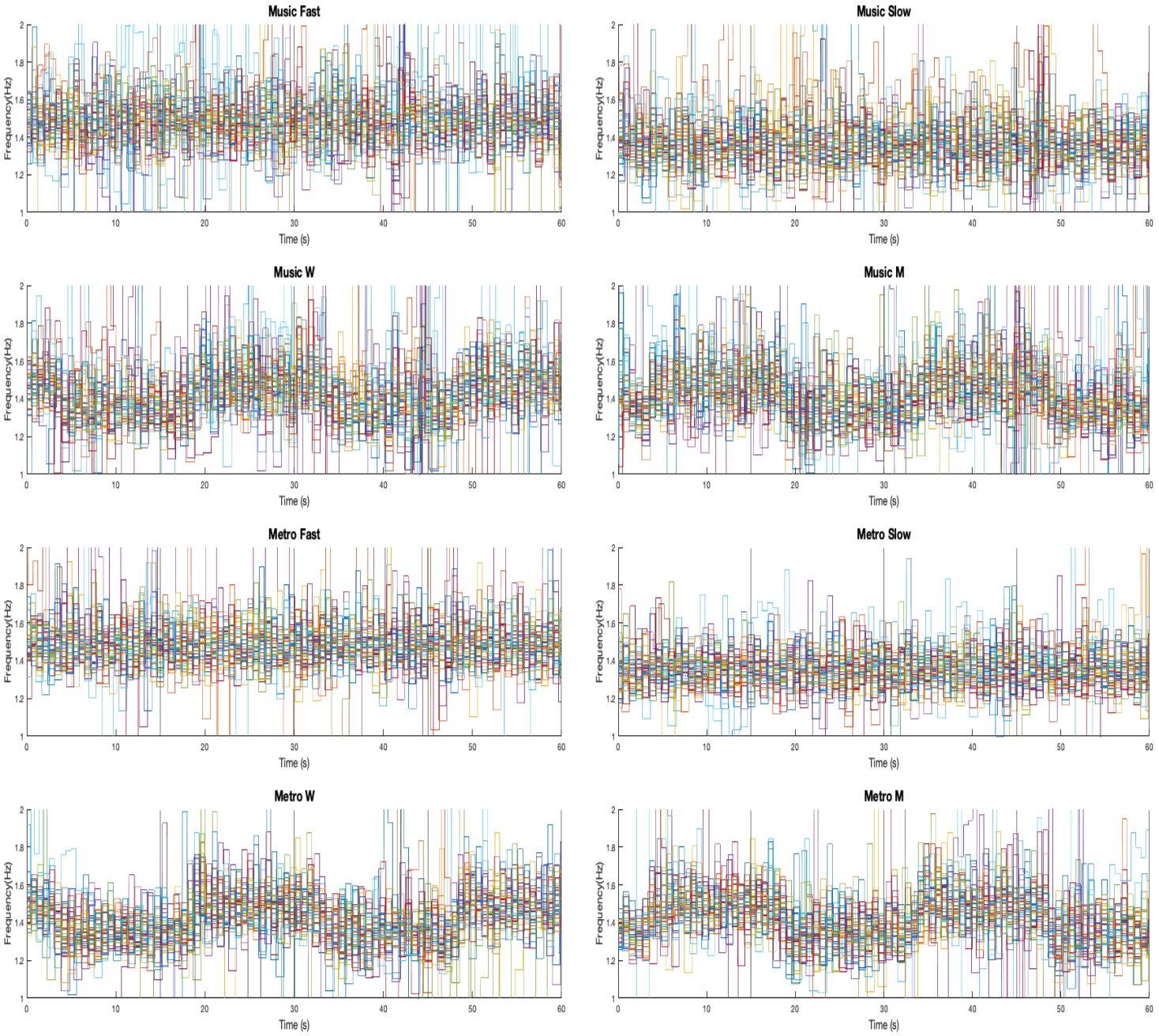


Figure 8. 11. Instantaneous frequency timeseries (finger-tapping).

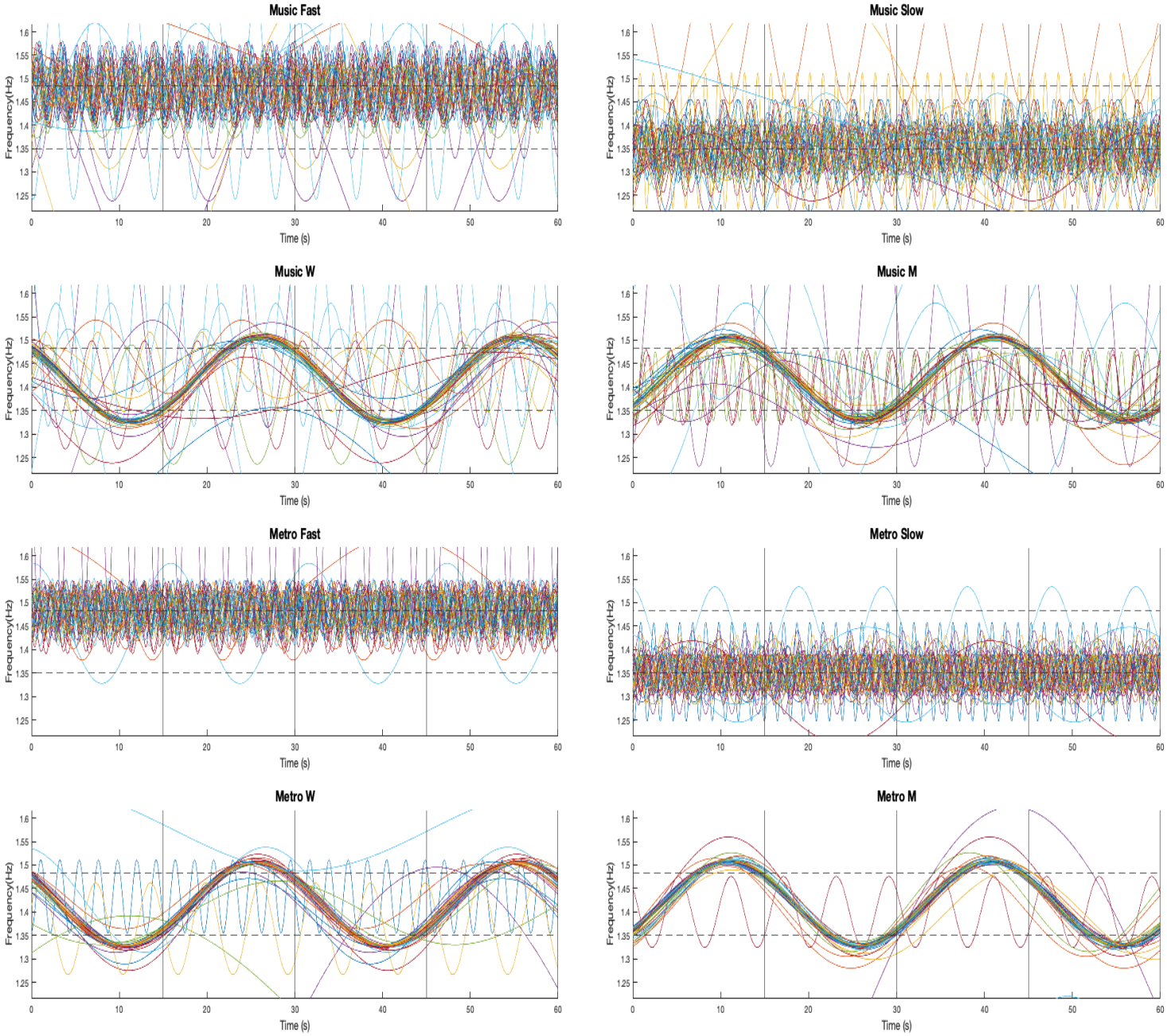


Figure 8. 12. Instantaneous frequency timeseries (sine curves fitted to the data).

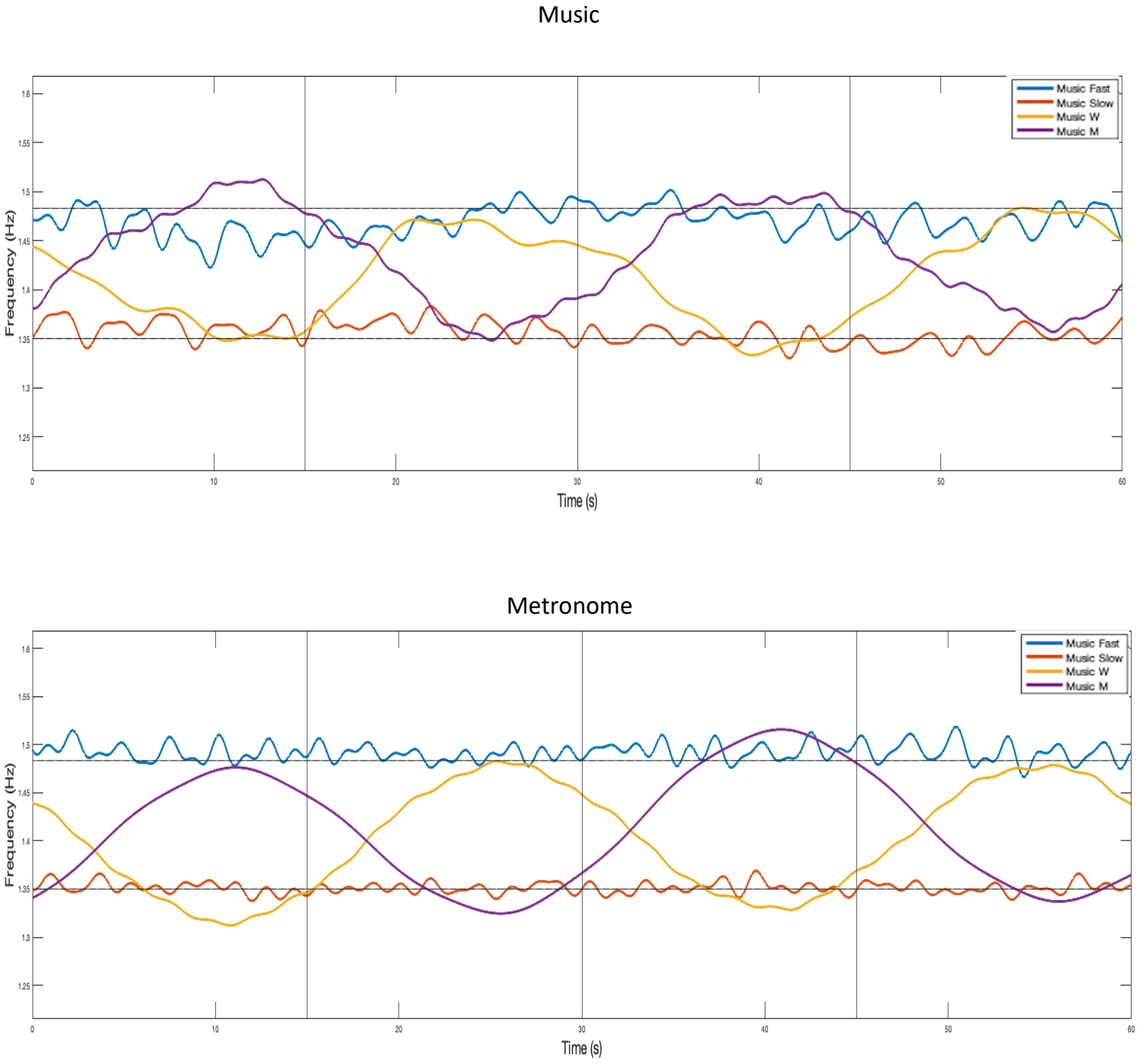


Figure 8. 13. Grand-average instantaneous frequency timeseries (sine curves fitted to the data).

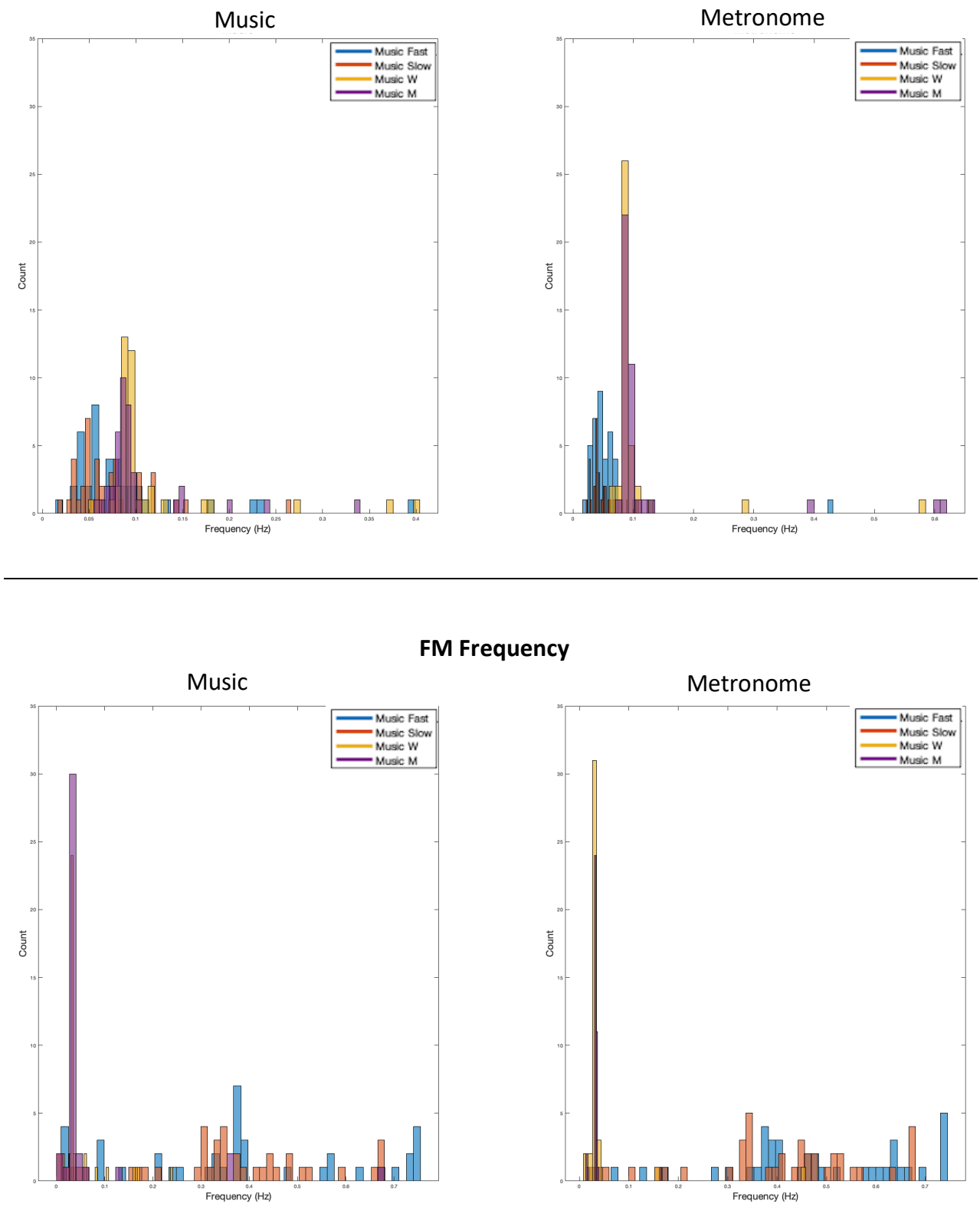


Figure 8. 14. Sine parameters distributions (sine curves fitted to the data).

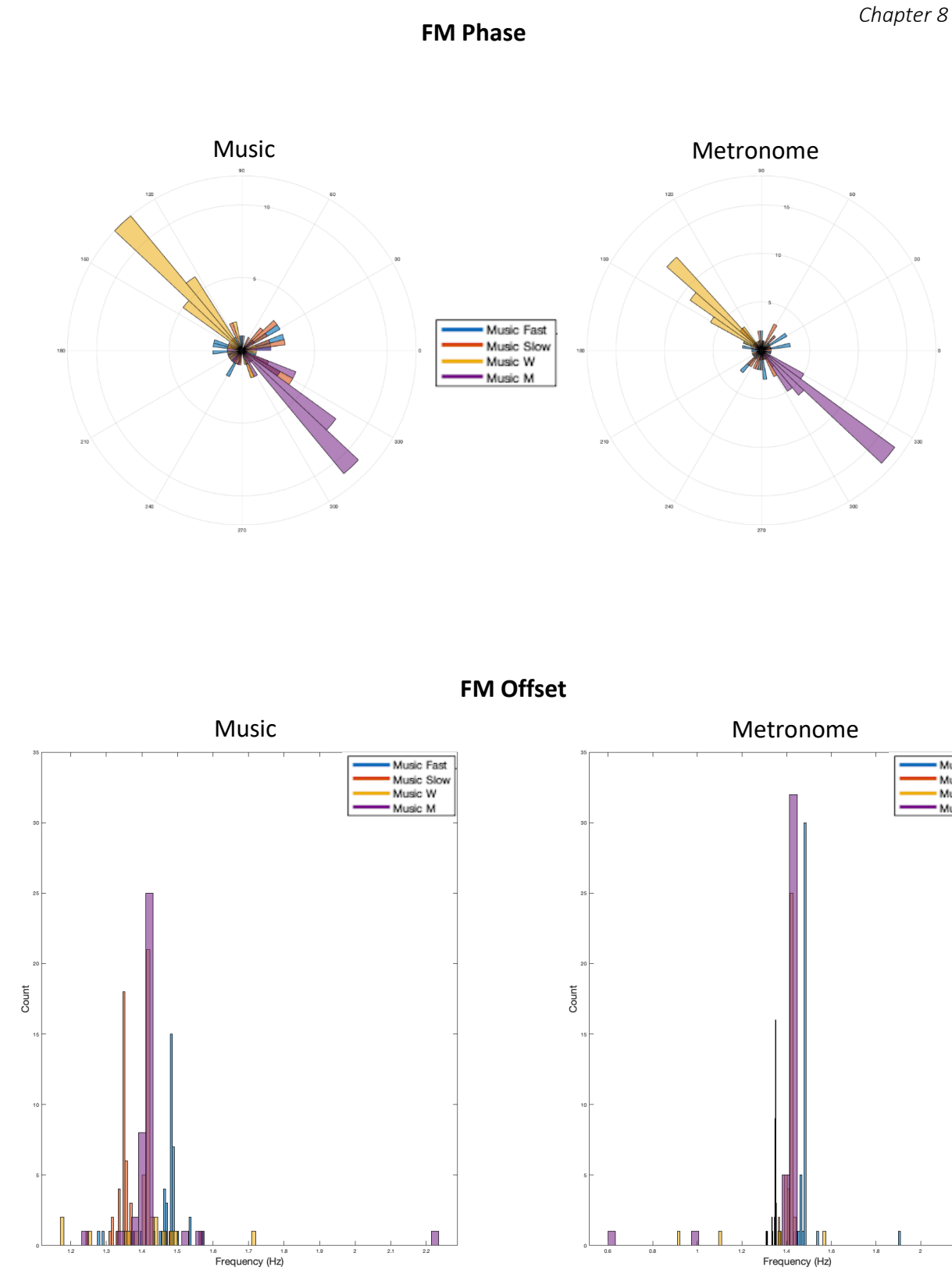


Figure 8. 15. Sine parameters distributions (sine curves fitted to the data).

		Music				Metronomes			
		Slow	**Fast**	**Change 'W'**	**Change 'M'**	**Slow**	**Fast**	**Change 'W'**	**Change 'M'**
Mean	**Offset**	1,472	1,357	1,41	1,433	1,492	1,351	1,398	1,408
	Amplitude	0,086	0,073	0,115	0,103	0,061	0,05	0,108	0,127
	Frequency	0,38	0,342	0,053	0,069	0,474	0,403	0,046	0,036
	Phase	2,753	2,85	2,761	4,755	2,83	3,183	2,454	5,236
Std	**Offset**	0,056	0,049	0,081	0,138	0,069	0,018	0,097	0,202
	Amplitude	0,069	0,045	0,073	0,052	0,064	0,022	0,084	0,124
	Frequency	0,238	0,186	0,049	0,124	0,189	0,174	0,07	0,023
	Phase	1,965	2,078	1,333	1,649	1,827	1,809	0,711	1,058

Table 8. 2. Descriptive statistics for sine parameters.

Approach #2: Event-related frequency adjustment (ERFA)

This second approach replicates the behavioral Event-Related Frequency Adjustment (ERFA) presented in Chapter 7. Instead of considering the data in its longitudinal format, here we identified the perturbation onsets and divided the time into 3-second post-perturbation time windows. To maximize the limited number of events, we aggregated the responses from slow-to-fast and fast-to-slow across the M and W conditions, resulting in a total of 4 responses per direction. For every participant, these 4 trials were averaged.

Despite the suboptimal signal-to-noise ratio resulting from the limited trials factored into the average, the responses consistently transition from one frequency level to the new one following the tempo change in both accelerating ('ERFA +') and decelerating ('ERFA -') conditions. These responses can be contrasted with the baseline levels of stable tempo conditions ('Fast' and 'Slow'). Given these observations, we identified the sigmoid function as a suitable candidate for modeling the transition with a few informative parameters. This is expressed in the following equation:

$$f(x) = \frac{L + (M - L)}{1 + e^{-b(x - c)}}$$

We emphasize that this function is optimal for capturing the transition from one stable tempo to another, in both directions of change. In the context of the perturbations, we propose the following interpretations for the parameters:

- range (M - L): given by the difference between the maximum and the minimum value, it indicates the magnitude of the adaptation to the step change (expected: 10% change with respect to the initial tempo);
- midpoint (c): it is the inflection point of the curve along the time x-axis, indicating how quickly the participant responds to the perturbation (providing a quantification of reaction time);
- slope (b): represents how quickly the curve transitions from one tempo to the other (the steeper the curve, the more abrupt the transition around the midpoint).

The parameters were estimated by optimizing the fit with the *sigmoid_fit()* function in Matlab, and the fitted curves are shown in *Figure 8.15*. The distributions of the parameters are depicted in *Figure 8.16*. We identified 3 participants deviating by more than 3 standard

deviations from the mean values of either midpoint or slope parameters in the ERFA – (decelerating). These were considered outliers and removed. The descriptive statistics of the sigmoid parameters' distributions are reported in *Table 8.3*.

The first notable difference from the plot of the sigmoid fits is that they exhibit a very sharp step. This is because the instantaneous frequency timeseries were produced by interpolating discrete data, and averaging just four trials was not sufficient to significantly smooth the curves. Due to the discrete nature of the data, the temporal resolution is compromised, which would be compensated for by averaging the responses to a larger number of events. However, we could still extract some valuable information from the estimated parameters.

The range fell within the expected range corresponding to a 10% step size: from 1.350 Hz to 1.483 Hz for 'ERFA +', and from 1.483 Hz to 1.350 Hz for 'ERFA –', indicating that participants were generally adapting to the tempo change in the expected direction. The transition midpoint exhibits more variability, suggesting that there might be significant differences across individuals and across stimulus types and the direction of the change. The slope is possibly the least informative parameter, arguably because most of the fits result in a curve very close to a step change. It's reasonable to expect that this issue could be mitigated by employing a system that tracks finger position at a stable sampling rate (like the motion capture discussed in Chapters 1 and 4), or by using a greater number of trials (like the task presented in Chapter 7). Unfortunately, both options are challenging to implement with the target population.

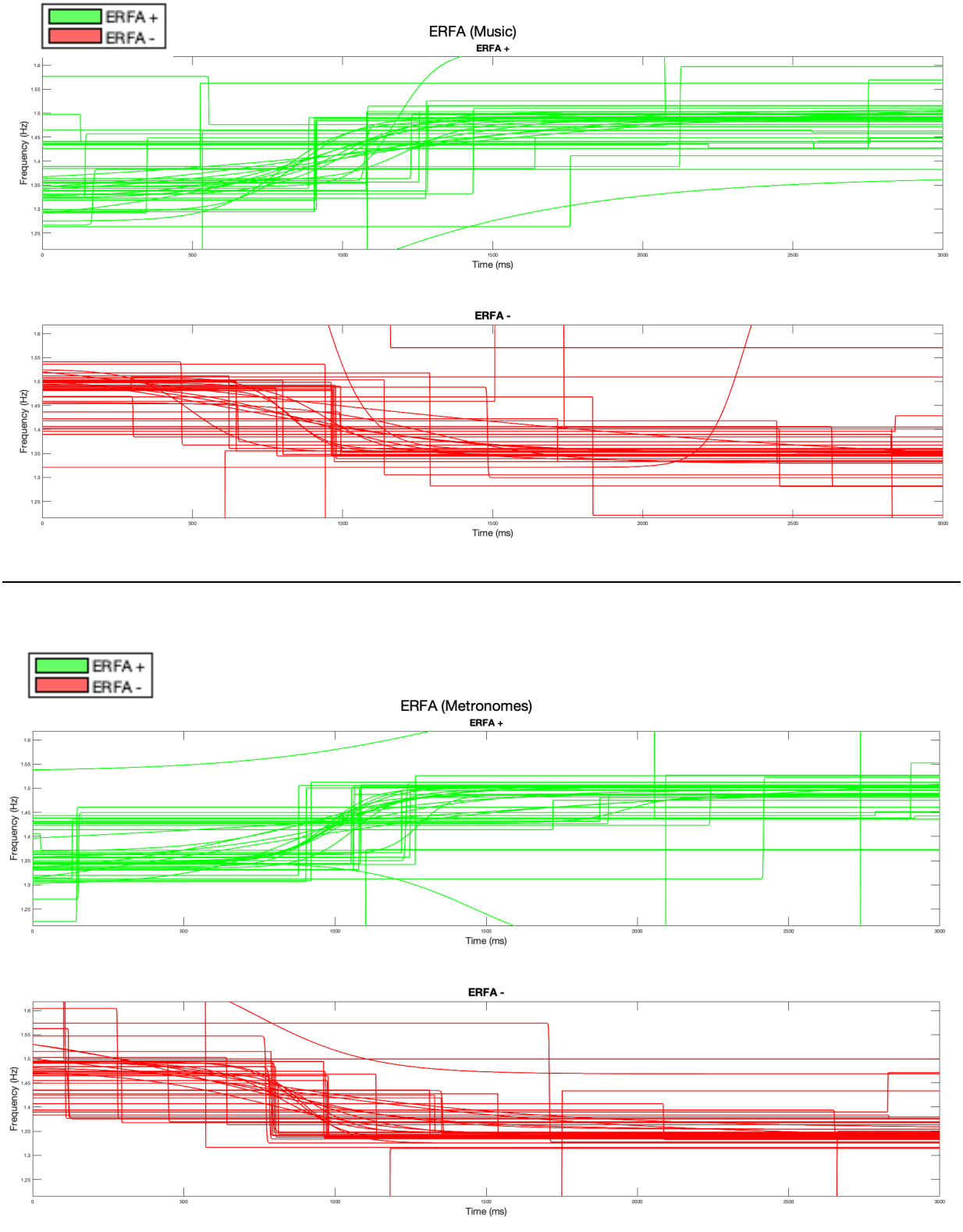


Figure 8. 16. Sigmoid fit to the empirical event-related frequency adjustments (ERFAs).

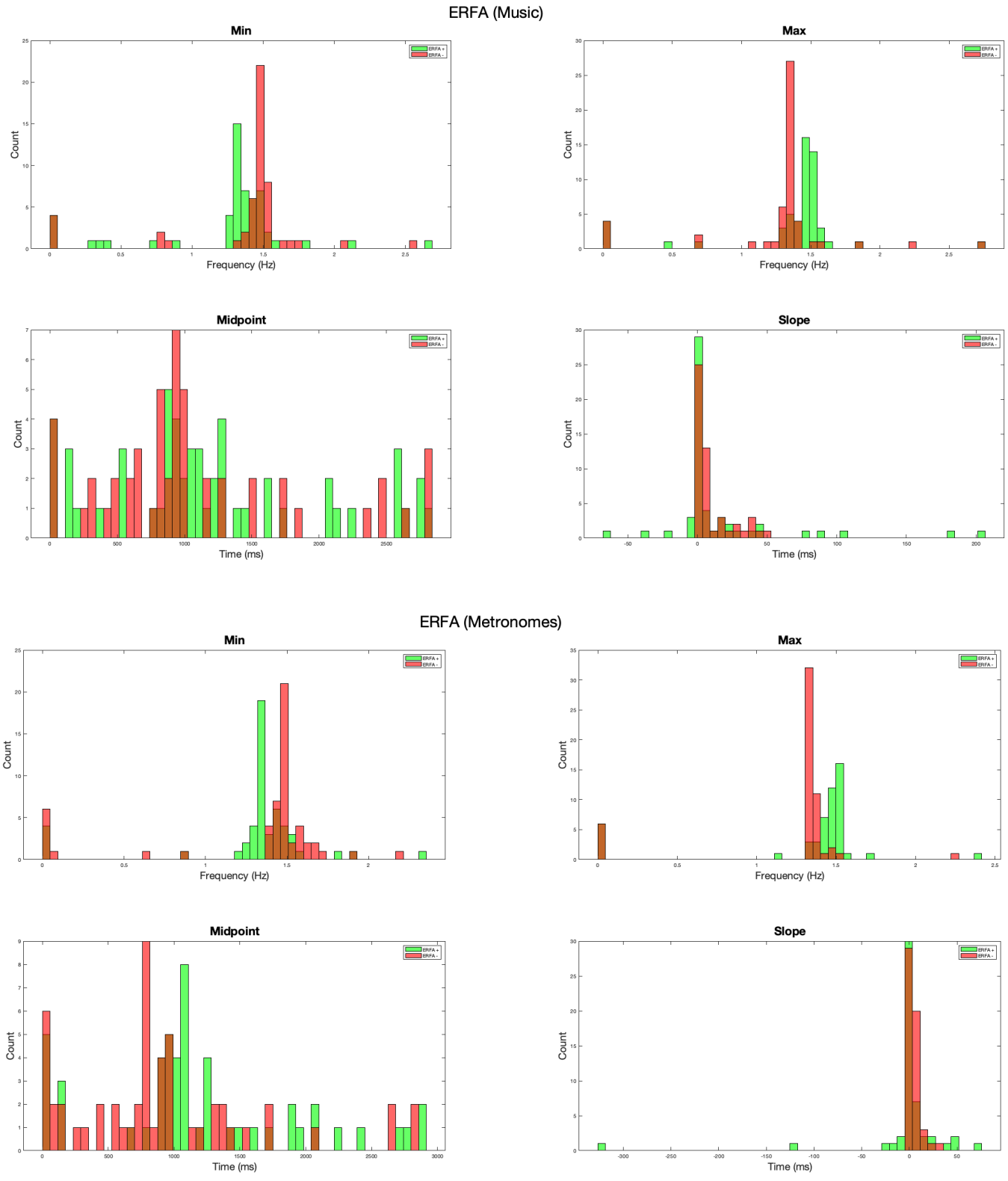


Figure 8. 17. Sigmoid parameters distributions.

Conclusion

The primary contribution of this chapter is to offer tools to enhance our understanding of the spatiotemporal dynamics of body motion and SMS, moving from the limitations encountered during data acquisition from NCD populations.

Musical interventions often revolve around SMS, due to the relative preservation of this functionality until the later stages of dementia. This is likely because cortical atrophy tends to spare motor structures in the early stages of the disease (Jacobsen et al., 2015), or because of compensatory rewiring reflecting attempts to compensate for their degeneration (Agosta et al., 2010; Ferreri et al., 2016). Either way, music is a rich communicative medium (Vuust et al., 2022) that involves the activation of a broad motor network (Grahn & Brett, 2007; Toiviainen, Burunat, Brattico, Vuust, & Alluri, 2020), with rhythm serving as the backbone supporting interaction with a music therapist. Given the potential of behavioral synchrony to mediate the rehabilitation of social and communicative skills (Cirelli, 2018; Tzanaki, 2022), it merits a thorough investigation of its underlying dynamics in order to optimize clinical interventions in the future.

This research implies a particular focus on how these dynamics change with disease progression and varying environmental contingencies, such as the level of presence of a music therapist and the nature of the auditory stimuli (Hobeika et al., 2021, 2022; von Schnehen et al., 2022). In the bigger picture of the research carried out at PSITEC, these methods serve to maximize and synthesize information in stage of data processing and modelling, providing informative outcome measures to complement central tendency and dispersion provided by circular statistics (Berens, 2009; Fisher, 1995).

For *dataset #1* (Ghilain et al., 2020), force plate data were used to represent the change in spatial patterns of body sway over time, during a rhythmic task with auditory stimuli in presence of a musician. QoM was computed as overall amount of movement, or changes in the center of gravity; RQA served to extract indicators of temporal dynamics of such motion, such as recurrence rate (RR) and determinism (DET); JRQA was performed to relate the temporal structure of motion to the mirroring of the musician's movements. A preliminary inspection of the results suggested a dissociation between the mechanisms underlying intentional SMS and the spontaneous motor activation of the whole body. While the original study reported an advantage for metronomic sequences in terms of SMS accuracy and consistency, our analysis on the force place data showed a major effect of ecological musical stimuli on motor engagement. Not only did we observe more movement elicited by music as quantified by QoM: we also reported that this movement was temporally organized to a greater extent by the musical stimulus (as quantified by

RQA), and by the interaction with the musician (as quantified by JRQA). It should be mentioned that the latter effect might have been mediated by the stimulus, working as a confounder which caused co-variance in the movement of the participant and the musician. The experimental design does not allow to draw final conclusions on the causal role of the interpersonal interaction.

Although the focus of this Chapter is methodological, we highlight that this opens a discussion on the interplay of different mechanisms participating to SMS. While it has been extensively discussed that auditory-motor interactions lead to synchronization, measuring different aspects of the phenomenon with a broad set of instruments is essential to complete the picture of the processes gravitating around pure SMS. Knowing how the population of interest interacts with different layers of the musical stimulus in an interactive context, it will be easier to leverage on the right musical features of musical stimulus for rehabilitation purposes. The relative ease of tapping along with isochronous metronomes may come at the expenses of the engagement of socioemotional and predictive mechanisms, which require emotional connotation and an optimal balance of temporal dynamics in the stimulus (Vuust et al., 2022).

In *dataset #2*, we employed two modelling approaches to capture the temporal dynamics of patients participating in a finger-tapping task with recurrent tempo changes. This task was designed to induce and measure adaptive behavior as patients recovered from synchronization errors (von Schnehen et al., *in preparation*). The first approach capitalized on the periodic nature of tempo changes in the macro temporal structure of the trials. We modelled the changes in the tapping frequency over time using sine functions, which allowed us to quantify the overall magnitude, consistency, and responsiveness of the adaptations to periodic perturbations. The second approach zoomed in on the transitions from one tempo to another, in critical time windows immediately following the perturbations. Despite the limitations imposed by the limited number of events available, this approach provided valuable information about the patients' responsiveness. Both approaches were based on the calculation of instantaneous frequency, which conveniently represents finger-tapping behavior in terms of tempo change. Each method presented here yields a single value which quantifies one particular aspect of the temporal dynamics of finger-tapping.

To the best of our knowledge, the techniques presented in this Chapter constitute the first application in the context of interactive musical tasks in clinical settings. The systematic tests on simulated, mock and real data offered a deeper understanding of the techniques, which can be adopted and adapted by other researchers in this domain. All outcome measures developed in this chapter are meant to be introduced as response variables in

the same statistical models previously developed by PSITEC for average-based measures or circular descriptive statistics, within the context of their respective experimental designs. We believe that our approach will contribute to a deeper understanding of this challenging line of research, maximizing the inference on behavioral dynamics during rhythmic interactions. The potential of music-based interventions aimed at improving the quality of life of elderly patients affected by NCD is considerable, and continues to motivate the development of methodologies for understanding their uniqueness and efficacy.

9

General discussion

Coordination dynamics

The aim of this dissertation was to unravel the dynamics at play when humans coordinate with each other, showing how individual rhythms can combine into organized forms of collective behavior. Conducted with a balance between individual and dyadic levels of analysis, we highlighted rhythm as a crucial element in interactions between humans and their environment. Considering such factor as a cardinal substratum for these interactions, we probed its role as a mediator and cornerstone for interpersonal coordination (P. E. Keller et al., 2014), allowing us to delve into underlying behavioral and neural mechanisms.

With the *drifting metronomes* paradigm at its backbone, we outlined an overarching methodological framework to induce the emergence of interpersonal coordination dynamics and facilitate their quantification, fulfilling our central objective. This approach enabled us to conduct a cohesive series of studies, in which we consistently observed that *interpersonal synchronization emerged spontaneously, against explicit instructions to ignore the partner*. Importantly, the instructions extended beyond the verbal domain: an intentional *competition* process was implemented via the incongruity between assigned rhythms, which participants actively attempted to maintain. However, we noted that a reciprocal *cooperation* process systematically began to dominate the interaction, effectively claiming a larger share of the coupling in the dyadic system, pushing the balance in favor of co-regulation at the expense of independency. This is significant, as we have not found other methods in existing literature that possess the properties of evenly spreading the coupling strength between the cue and the partner, controlling for intentional synchronization, and implementing an authentically competitive nature (for a detailed discussion on the properties of the drifting metronomes, see Chapter 1).

The primary focus here is the principle of an *attractor*, which operates as a dynamical structure latent to the interaction by driving the system towards a stable state (Marsh et al., 2009; R. C. Schmidt et al., 1990; R. C. Schmidt & O'Brien, 1997; Tognoli & Kelso, 2014; Tognoli et al., 2020). Instead of merely inferring this construct from observed behavior, we showed how the force of this attractor interacts with the intentional actions of the individual entities within the dyad. To put it another way, even when at odds with deliberate efforts to avoid any mutual influence, *humans naturally show a tendency towards interpersonal coordination*. Importantly, it was observed that dyads did not maintain a synchronized state at all times. When applying our analytical tools on a more granular level, it became evident that *the attractor bounded the interaction according to consistent dynamics*.

This is arguably the most interesting level of analysis, but what are exactly the dynamics we are discussing here? We specifically refer to a form of systemic behavior where integrative and segregative tendencies coexist among individuals (Ole Adrian Heggli, Konvalinka, Kringelbach, et al., 2021), giving rise to a temporally structured pattern in which periods of dwells in a coordinated state are broken by instances of divergence, or escape behaviors. This property is referred to as *metastability*, a concept widely used in the literature to point at a latent dynamical structure of the system, which facilitates both permanence within and transitions across states (Ole Adrian Heggli, Konvalinka, Kringelbach, et al., 2021; J. A. S. Kelso, 1995; J. A. S. Kelso, DelColle, & Schöner, 2018; J. A. S. Kelso & Tognoli, 2009; Tognoli & Kelso, 2014). The divergence of trajectories implemented in the drifting metronomes was explicitly designed as a scanning procedure, a tool for systematically exploring the metastable structure latent to the system. The layout of this hidden construct revealed an *attractor landscape* (Schöner et al., 1992; Tuller & Kelso, 1989; Yamanishi et al., 1980), which covers a variety of states that promote either one of the two opposing tendencies. Using the recurrence score as a proxy measure for the coupling between the partners (Marwan et al., 2007), we showed that dyads tend to oscillate between cooperation and competition as a function of the attractor landscape. Specifically, while transitioning over the in-phase region resulted in a long-lasting attraction on a stable collective state, the proximity anti-phase point facilitated decoupling among the partners and the pursuing of independent trajectories.

Without a doubt, the most solid finding of this entire dissertation is the empirical layout of this attractor landscape, replicated across the three studies presented in Chapters 2-4. All these studies tested conditions of face-to-face, visually mediated coupling, providing the base case for dyadic interactions (Richardson, Marsh, Isenhower, et al., 2007; Richardson et al., 2005; R. C. Schmidt et al., 1990; R. C. Schmidt & O'Brien, 1997). The consistent findings not only underscore the inherent human tendency towards coordination, revealing the dynamics governing the balance between cooperative and competitive tendencies within a dyad: it also validates our methodological approach. Our expansion on earlier literature strengthens the understanding of the attractor as a latent construct (Tognoli & Kelso, 2014), its landscape (Schöner et al., 1992; Tuller & Kelso, 1989; Yamanishi et al., 1980), and, most importantly, extends this concept to the interpersonal domain (Marsh et al., 2006, 2009; R. C. Schmidt & Turvey, 1994). As Scott Kelso stated in his influential book *Dynamic Patterns*, "one's belief in the scientific method is enhanced when a given set of findings is confirmed or replicated by others, especially if it happens to be in a slightly different paradigm" (Kelso, 1995, p. 104).

Informational coupling

Having covered *coordination dynamics*, we now pivot to the other key aspect of our discussion: *informational coupling*. Noting that temporal coordination between individuals is mediated by information exchange via sensory channels (R. C. Schmidt et al., 1990), we highlighted the multifaceted nature of its basis. Given that this exchange establishes interpersonal action-perception loops, which lay the groundwork for coordinated behavior (Phillips-Silver et al., 2010; Tognoli et al., 2020), we put forward the working hypothesis that *the attractor layout latent to the dyadic system is dependent on how individuals exchange, access, and process mutual information*. This hypothesis was tested throughout a series of studies where different dimensions of informational basis were manipulated as independent variables, to assess their effect on the dyadic attractor layout. Starting from the most discernible source of variability, namely which sensory channel mediates the coupling, and building on subsidiary dimensions, each study was designed to tackle specific research questions individually.

Question *Chapter 2:* *are attractor dynamics dependent on sensory modality?*

Here, we explored the role of sensory modalities in mediating informational coupling, within the context of spontaneous dyadic entrainment, comparing the attractor landscape returned by the drifting metronomes procedure in visual and auditory coupling. We found that visually-mediated entrainment was governed by a highly consistent layout across dyads, expressing the opposing tendencies described above. In contrast, entrainment mediated by auditory coupling led to more variable profiles, with strong tendency to synchronize with the partner at all times, regardless of the states being visited. This suggests that the *coordination dynamics differ depending on which perceptual system samples information from the other individual*. While visual coupling loses traction when the dyad is far enough from the influence of the in-phase attractor, auditory coupling consistently leads to the convergence of individual rhythms.

Ultimately, the structure of the interaction was clearly governed by dynamics that varied depending on the sensory modality in play. However, the study left an open question. Due to the design of the stimuli, we could not rule out that differences could be explained by the availability of kinematic information. To clarify, while visual coupling enabled participants to continuously sample the entire phasing of the partner's movement, with

auditory coupling they had to rely on a discrete and irregular sampling of the partner's taps at the moment of impact. This made the experimental design unbalanced with respect to this ancillary dimension, calling for additional investigation to control for the role of kinematic information.

Question *Chapter 3:* *are attractor dynamics dependent on the access to kinematic information?*

Building on the previous study, we put forward the hypothesis that differences across modalities would fade, or at least attenuate, when balancing the access to movement kinematics across visual and auditory coupling. Arguing that the physical property of the stimulus and the perceptual system accessing such information constitute two distinct dimensions of informational coupling, we designed an experiment to assess their relative contributions to the interaction. This was achieved by expanding on previous experimental conditions, augmenting auditory coupling via continuous sonification of the partners' motion, and discretizing visual coupling via its occlusion.

Our results showed that *attractor dynamics are invariant to the continuous or discrete access to kinematic information*, thereby confirming their dependence on sensory modality. This led us to conclude that differences are not a matter of access to kinematics, but should be explained based on the unique interface of each sensory systems with motor and timing systems (Comstock et al., 2018). We put forward that, due to the temporal specialization of the auditory system, the attraction was so strong that it did not allow individuals to decouple at any time. On the contrary, while dominant in proximity of the cooperation attractor point, visual coupling enabled individuals to segregate the information and pursue individual rhythms, taking advantage of the competition attractor promoting divergent behavior.

This work suffers from the limitation of relying on only one sonification strategy. Future work should explore alternative approaches which may be more effective in conveying kinematic information in the auditory domain, which may result in a more effective assimilation of visually- and auditorily- mediated dynamics.

Question *Chapter 4:* *are attractor dynamics modulated by embodied perspective taking?*

This work was explicitly situated at the intersection between fundamental research on interpersonal coordination, and the deployment of technology for guiding and supporting human rhythmic interactions. A critical aspect of virtual reality (VR) technologies is the

embodiment of the user, whose sensorimotor processes can be modulated by these technologies at various levels of immersion (Maes et al., *under review*). By implementing a body-swap illusion (Petkova & Ehrsson, 2008; Petkova, Khoshnevis, et al., 2011), we explored the potential of manipulating visual perspectives for modulating synchronization during joint actions. The investigation was circumscribed to the visual domain, of which perspective was treated as subsidiary dimension.

On the fundamental side, we observed that dyadic coordination dynamics are subject to the manipulation of visual perspective. Participants in 1st person coupling were less effective at maintaining decoupled trajectories around the competition attractor, as compared to those in 2nd person coupling. These results suggest that visual perspective influences coordination dynamics in dyadic interactions, engaging error-correction mechanisms in individual brains (Maes et al., 2015; Mazzoni & Krakauer, 2006; Smith et al., 2006; Wagner & Smith, 2008) as they integrate the partner's hand into their body representation (Maister et al., 2015; Tajadura-Jiménez et al., 2012; Tsakiris, 2010, 2017). On the application side, we demonstrated that VR technology can be used to influence social interactions at the lowest level of coordination dynamics, thereby supporting joint action through enhanced interpersonal synchronization. We propose that a technology informed by principles of body-swapping has considerable potential for facilitating interpersonal coordination across a wide range of applications, including motor training, sports, musical education, and rehabilitation protocols.

Formalizing a framework

Taking all this evidence together, we hereby propose to capture in an equation the *cooperation / competition balance*, as it was investigated across studies with the drifting metronomes paradigm. The following is not meant to be a rigorous mathematical formula, but rather a high-level, descriptive expression to synthesize the concept of this series of studies and integrate the experimental designs in an organic fashion. It should be stressed that its validity is limited to the drifting metronomes paradigm in dyads, and should not be generalized to ecological interactions in its current form.

$$\Phi = C * \left[\left(\beta_0 + \beta_1 \Psi + \beta_2 \Psi^2 \right) : \begin{vmatrix} modality \\ kinematics \\ perspective \end{vmatrix} \right] + \mathcal{E}$$

Where:

- Φ is the *order parameter* of the system, namely a variable which captures the temporal relations among individuals within the dyad. In our studies, a recurrence score (Marwan et al., 2007) was used to quantify the degree of temporal coordination between the partners, and interpreted as an indicator for coupling strength.

- *C* indicates the presence of *coupling*, manipulated in the studies as a dichotomic independent variable. It can assume 0 and 1 values, reflecting OFF and ON states, respectively. Note that, when set to 0, the coupling is disabled and the whole term inside square brackets equals 0[1].

- Ψ is the *control parameter*, operationalized here as the relative phase between the metronomes and used to drive the dyad through the space of coordinative states. It maps onto the time dimension of the trials. Due to its circularity, the in-phase point is visited at the extremes of the trials, whereas the anti-phase point is in the middle. This symmetrical structure motivated polynomial modelling in the coefficients β.

- β are the coefficients of the 2nd order polynomial used to model the changes of the *order parameter* as a function of the *control parameter*. β_0 is the intercept and represents the average level of coordination between individuals; β_1 is the linear term and affects the shift of the parabolic modulation vertex and relates to the hysteresis of the system; β_2 is the parabolic term and expresses the depth of the modulation as a function of the attractor landscape.

- | | contains the factors tested in each individual study, which have both main effects on the order parameter and interaction effects with the attractor dynamics expressed in the polynomial model. The factors can be arranged in a nested hierarchical structure, which are not illustrated here for simplicity. For example, we postulated *modality* and *kinematics* to be orthogonal dimensions, and manipulated them as such in our

[1] The operator ' * ' is used here as arithmetical operator. The operator ' : ' is reserved to indicate interaction effects in a statistical model.

experimental design. Conversely, *perspective* was operationalized within the visual domain and therefore considered subordinate to *modality*. This is the level where future studies can bring their contribution in this model by including new factors in their experimental designs. Expanding this set of dimensions would then allow to better outline their relationships.

- $\mathcal{E}$ is the error term, explained by variability inherent to the individuals' motion. When *C* is set to zero, this term entirely accounts for spurious patterns observed in Φ.

As final remark on our behavioral findings, we want to mention where they find their place as a whole with respect to the major theories currently debating interpersonal synchrony (Palmer & Demos, 2022). Since none of the studies were designed to explicitly test predictive accounts against dynamical systems theory, a balanced discussion from both angles was provided in the discussion of every Chapter.

Nevertheless, we should mention that the findings of Chapters 2-3 are quite compatible with a parsimonious oscillatory framework, where perceptual and motor oscillators mutually interact (Ole Adrian Heggli et al., 2019) and a coupling term is modulated by the factors at play and by the attractor layout. Chapter 4, conversely, is hardly explainable without involving some form of internal models, since the effects of our body-swap manipulation seem to be necessarily mediated by a transient change in the representation of the bodily self (Tsakiris, 2010) or by a modulation of internal models for motor control (S. J. Blakemore et al., 1999; S.-J. Blakemore, 2017; Wolpert et al., 2003).

We believe that the explanation provided by the dynamical system theory is the preferred solution in terms of parsimony, but reaches a limit when dealing with a level of cognition that calls for more elaborated central representations. However, we do not see these theoretical contenders as mutually exclusive. Just like fundamental mechanisms at the lowest level of basic dyadic interaction can serve more high-level functions underpinning complex forms (Knoblich & Sebanz, 2008), informational coupling may be subject to a hierarchy of nested cognitive functions. In building up this first corpus of studies, we explored the dimensions of this multifaced construct in increasing complexity, probing the explanatory extent of the theories at every step.

Neural dynamics

A second series of works was dedicated to the investigation of the neural underpinnings of human rhythmic interactions. By means of individual and dual EEG recordings, we aimed at unveiling different mechanisms by which means the brain tracks the dynamics of environmental and human rhythms, enabling overt synchronization behavior. Building on the concepts of cross-frequency coupling and entrainment, considered essential forms of interaction between neural and environmental oscillatory processes (Allen et al., 2011; Canolty et al., 2006; Michael X. Cohen, 2008; Lakatos et al., 2019), we targeted neural dynamics at the level of interaction between internal and external rhythms.

Intrinsic to the organization of brain physiology is its adaptability, predisposing the brain to dynamically interact with changing environments. This suggests that adaptive responses necessitate proficiency in tracking these changes to facilitate flexible behavior (Criscuolo et al., 2022). Such adaptability is thought to be underpinned by the coupling of internal and environmental rhythms (Large & Jones, 1999), and by predictions encoded in the temporal scaffolding of frequency-specific oscillatory dynamics (Nobre & van Ede, 2018).

Proving the causality of neural processes by means of non-invasive electrophysiological recordings is an extremely hard endeavor (Novembre & Iannetti, 2021). While none of these studies achieved this feat, we provided a solid corpus of correlational evidence, along with a wide range of innovative techniques whose design was carefully tailored to capture specific dynamics elicited during the task. In answering the following research questions, we contribute to further grounding the discussion of brain theories on the underlying physiological processes (Palmer & Demos, 2022).

Question Chapter 5: *how are beta oscillations involved in interpersonal synchrony?*

Hyperscanning EEG recordings were performed in the study presented in Chapter 2. In analyzing them, we proposed a method to overcome a criticality of traditional measures based on interbrain synchrony, which typically fail to account for the elements of action and perception that lead to its observation. We argued that, in order to foster the interpretation of interbrain dynamics, one should explicitly model the involvement of the effectors as bodily oscillators in the setting. Because the effectors in action represent the observables accessed by two independent nervous systems via perceptual coupling (R. C. Schmidt et al., 1990), they necessarily mediate brain dynamics relevant to the interaction.

By extracting an oscillatory component in the beta range (~20 Hz) and quantifying its modulation as a function of finger-tapping cycles, we found that this was significantly modulated by both self-generated and other-generated movements. In conditions where partners perceived each other, we observed periodic fluctuations of beta power corresponding to reciprocal movement cycles. This modulation occurred in both visually and auditorily coupled conditions and was accompanied by recurrent periods of dyadic synchronized behavior.

Our study provides evidence that the brain employs a shared oscillatory mechanism to orchestrate self-generated rhythmic movements and to track those produced by a partner. This dual function is proposed to be simultaneously represented in an oscillatory expression of the common coding theory (W. Prinz, 1990; Wolfgang Prinz, 2013), a principle understood as crucial for facilitating temporal coordination among individuals (P. E. Keller et al., 2014). These findings bridge two bodies of evidence: one supporting the general role of beta oscillations in encoding rhythmic sequences (T. Fujioka et al., 2009; Takako Fujioka et al., 2015, 2012), and another substantiating its specific contribution in encoding movement timing and kinematics (Nijhuis et al., 2021; Novembre et al., 2017; G. Zhou et al., 2016). Thus, *mutual modulation of beta power* emerges as a potential mechanism underpinning the real-time integration of self-generated and externally observed actions.

Question *Chapter 6:* *is the stability of entrained oscillations an index for auditory-motor coupling?*

Here, we sought to further explore the interaction between auditory and motor systems and their influence on rhythmic behavior (Morillon et al., 2019, 2014; Morillon & Baillet, 2017). This interaction was investigated at the level of entrainment, as a putative mechanism of alignment between endogenous brain oscillations and external auditory rhythms. This understanding is particularly beneficial in the realm of neurological rehabilitation, especially in improving locomotive capabilities (Ashoori et al., 2015; de Dreu et al., 2012; Moumdjian et al., 2018; Moumdjian, Moens, Maes, Van Nieuwenhoven, et al., 2019). For this reason, this work was originally motivated by the need for an outcome measure to quantify the stability of auditory-motor coupling during a rhythmic interaction with auditory cues, complementing behavioral measures traditionally used in clinical research (Moumdjian et al., 2018).

Our study involved young healthy participants engaging in a basic finger-tapping task, who were instructed to synchronize with an isochronous auditory metronome. We employed

a novel computational method to extract a neural measure from the EEG signals recorded during this task. Our focus was to identify and isolate an oscillatory component from the broadband brain activity, maximally attuned to the stimulation frequency (Michael X. Cohen & Gulbinaite, 2017), and quantify the degree of fluctuation in this frequency over time (Michael X. Cohen, 2014). Our procedure resulted in a *stability index*. This measure is interpreted as an indicator of the brain's ability to maintain stable entrained neural oscillations in response to a steady auditory rhythm.

We validated our neural measure by correlating it with several indicators of the participants' performance in the synchronization task. We found that the smaller the stability index (with 0 indicating a perfectly stable oscillation), the better the participants performed in terms of behavioral accuracy and consistency. Our findings underscore a correlation between the stability of attuned neural oscillations and effective behavioral synchronization, indicating its potential as a proxy measure for neural auditory-motor coupling. This could have practical implications for neurological rehabilitation, such as using the stability index as an outcome measure to evaluate the effectiveness of rehabilitation protocols in populations with compromised sensorimotor abilities due to an underlying pathological condition.

However, in terms of fundamental insights on neural entrainment, the study was limited by the lack of systematic manipulation of stimuli to elicit predictable dynamics. Given these limitations, it was necessary to further investigate this mechanism under more controlled conditions, which brings us to the next study.

Question *Chapter 7:* *does neural entrainment underpin sensorimotor synchronization to dynamic rhythmic stimuli?*

Despite a surge in empirical studies and discussions (Lakatos et al., 2019), validating the premise of neural entrainment in humans has been considerably difficult. While the concept is theoretically straightforward, capturing its dynamic underpinnings through non-invasive electrophysiology has been elusive (Haegens & Zion Golumbic, 2018; Novembre & Iannetti, 2018; Rajendran & Schnupp, 2019). Our research method evolved from the approach presented in Chapter 6, and involved the manipulation of dynamic rhythmic stimuli as essential means to induce and observe the hypothesized mechanism in action. The design of spatial filters (Michael X. Cohen, 2022; Michael X. Cohen & Gulbinaite, 2017) enabled us to isolate perceptual and sensorimotor oscillatory components attuned to the stimulation frequency from multivariate EEG signals. Notably, by implementing perturbations in the rhythmic stimuli, both types of components

demonstrated their adaptability by adjusting their frequency in response, aligning their oscillations over time to match the changing stimulus dynamics. This controlled approach allowed us to draw inferences about the nature of neural entrainment, as we observed sensorimotor processing to enhance the entrained response, thereby strengthening the notion that the active engagement of the motor system contributes to processing rhythmic stimuli (Haegens & Zion Golumbic, 2018; Morillon & Baillet, 2017; Morillon et al., 2015, 2014; Rimmele et al., 2018).

With the novel use of *event-related frequency adjustment* (*ERFA*), we have progressed towards a more precise quantification of neural entrainment via non-invasive electrophysiology. Our findings lend compelling support to the idea that this putative mechanism effectively underlies overt sensorimotor synchronization. By moving beyond amplitude-based quantification of neural responses (Lenc et al., 2018; Nozaradan et al., 2011, 2015), we explicitly focused on the phase dynamics of oscillatory components, contributing to the current discourse in the field (Rajendran & Schnupp, 2019).

Our empirical findings are further supported by a series of computational simulations, where we contrasted our oscillatory model, grounded in endogenous neural oscillations, against an alternative model of evoked responses which does not assume ongoing oscillatory activity (Novembre & Iannetti, 2018). Our results favor the oscillatory model, suggesting that entrainment is not merely an evoked response passively tracking changes in stimulation rate, but indeed a process of active alignment of neural oscillations to external rhythmic stimuli.

In sum, we've made considerable strides in addressing one of the most pressing challenges in human neuroscience (Haegens & Zion Golumbic, 2018). While not able to definitely rule out the contribution of evoked responses to our measure, our results provide compelling evidence for neural entrainment as an underlying mechanism for overt behavioral alignment to dynamic stimuli, marking another significant step forward in the journey to understand the underpinnings of sensorimotor synchronization.

A toolkit for analyzing rhythmic interactions in vulnerable populations

On a dedicated Chapter, we circumscribed a set of methodological contributions in the framework of a collaboration between *IPEM - Institute for systematic musicology* (Ghent University, BE) and *PSITEC Laboratory* (University of Lille, FR). Here, we navigated the challenges of conducting research within vulnerable populations suffering from neurocognitive disorders (NCDs). These challenges often limit data acquisition, but the work detailed here has strived to overcome these hurdles and maximize the inferences on sensorimotor synchronization in these patients, thus making a significant contribution to the primary objectives of this thesis.

Question *Chapter 8:* *how can we facilitate the investigation on SMS in healthy and pathological aging?*

The study of SMS is central to understand the effectiveness of music-based interventions on patients suffering from neurocognitive disorders (NCDs) (von Schnehen et al., 2022). In this endeavor, it is of outmost relevance to understand how motor, expressive, and empathic processes are tied to this basic function during musical interactions (Hobeika et al., 2021, 2022). Despite the potential benefits, working with NCD patients poses some unique challenges. For instance, the application of invasive recording devices may not be well tolerated and cause physical discomfort or fear responses. These patients also have limited tolerance for long, repetitive experimental sessions that require continuous movement, influencing the data collected and making it less generalizable.

To overcome these challenges, the experimental procedures for this population need to prioritize non-invasiveness, brevity, minimal repetition, engagement, and human support in a familiar environment. The chapter proposed to analyze these datasets using the principles established in the dissertation, providing a deeper understanding of spatiotemporal dynamics from musical interactions tailored to NCD patients. Ultimately, the tools developed here attempted to maximize the use of available timeseries data at the stage of processing and modelling, thus seeking to enhance our understanding of how healthy and pathologically aging individuals respond to musical interventions.

The expansion of dynamical principles previously applied to young healthy participants proved to be a precious resource in this domain. In particular, source-separation methods applied to force plate data (Desmet et al., 2017), in combination with recurrence

quantification analysis (RQA) (Demos & Chaffin, 2017; Demos et al., 2018; Marwan et al., 2007, 2002) allowed to move beyond the mere quantification of motion and extract complexity measures related to the temporal structure of body motion. Preliminary data inspection suggested a dissociation between the mechanisms underlying intentional SMS and the spontaneous activation of the whole body. Specifically, it was observed that ecological musical stimuli had a major impact on motor engagement, in contrast with previous findings of improved SMS with metronomes (Ghilain et al., 2020). This motivates to further delve into the interplay between different socio-emotional mechanism and SMS, backed by the newly introduced analytical tools. Moreover, new approaches built on the concept of ERFA allowed to model the temporal dynamics of adaptive synchronization behavior of patients, which is a crucial aspect of SMS (von Schnehen et al., *in preparation*). This methodology unraveled the patients' responsiveness to tempo changes, enabling to parametrize several dimensions of the temporal dynamics underlying error correction.

Through these findings, we contributed to a deeper understanding of behavioral dynamics during musical interactions, paving the way for the development of music-based interventions to improve the quality of life of elderly patients affected by NCD. The proposed analysis toolkit is expected to facilitate further developments of this delicate line of research.

Future directions

Why invest such substantial effort in refining an experimental paradigm? Why place such emphasis on it? The answers lie in our overarching aim throughout this dissertation: to provide a methodological framework and a comprehensive corpus of both behavioral and neurophysiological evidence. By doing so, we hoped to lay a solid foundation for an expanding line of research in the domain of interpersonal coordination, a platform from which fellow researchers could design their experiments and pursue their unique research questions. Furthermore, the analytical tools we introduced or refined, including innovative EEG methods such as mutual beta modulation, stability index, and ERFA, are not fixed in their current form. We hope to see these tools adopted, adapted, improved and built upon, facilitating a dynamic evolution that aligns with the demands of this field.

The strength of this work lies not only in its results but also in its potential for growth and expansion. By acknowledging both its merits and limitations, we have attempted to establish a pragmatic, transparent, and reproducible starting point for future investigations. The limitations, in particular, should be viewed as open doors leading to

still unexplored paths. We hope they will inspire future studies, driven by new questions and hypotheses. We therefore conclude this dissertation by addressing some relevant lines of inquiry and possible directions that future research may take, illustrating the broader scope and potential of the perspective we outlined.

Expanding experimental designs

Through a first series studies, we managed to disentangle some fundamental dimensions of perceptual coupling and assess their effects on the dynamics of interpersonal synchronization. However, this work only scratches the surface, with many aspects yet to be explored. An immediate extension of our work involves identifying additional relevant dimensions, and directly build on our methodological framework by manipulating new independent variables of interest in *within-subjects* designs. Alternatively, different populations of interest can be tested under the same conditions using a *between-subject* design. We suggest some straightforward extensions for future studies below:

- *Spatial dimension* and *classes of movement*. We point out that the spatial dimension was not manipulated in our dyadic studies but controlled by constraining the finger-tapping movements along the vertical axis. It would be of interest to introduce the spatial dimension as an experimental variable in the design, by manipulating spatial congruency in the movements along orthogonal axes. Seminal work by Kilner et al. (J. M. Kilner et al., 2003) shows that, when movements are incongruent, mirroring another individual causes motor interference and spontaneous spatial alignment. It would be of interest to investigate how congruency interacts with temporal dynamics, and specifically whether cooperation and competition tendencies are facilitated by congruent and incongruent movements, respectively.
  Within the same paradigm, it would also be possible to enable different classes of movements and effectors, increasing the ecological validity of our findings beyond the limits of a finger-tapping task.

- *Personal variables*. Another straightforward extension would consist of including *between-subject* manipulations, to investigate the effect of personal variables on coordination dynamics. Whilst in our works musical expertise was controlled and matched across participants, one may explicitly include in the design a group of experts in different musical instruments and genres. For instance, dyads of expert drummers are capable of segregating information from the other and carrying out

synchronization tasks with minimal mutual influence (Ole A. Heggli et al., 2019), which makes them an ideal model of sensorimotor experts[2]. Alternatively, different configurations of dyads can be formed according to social variables, whereas in our studies dyads were gender balanced and the participant's ethnicity was randomly sampled. This could turn out to be an informative means to test attractor dynamics as implicit measure of pro-social affiliation (Marsh et al., 2009; Tzanaki, 2022).

- *Pathological models*. Additional between-subjects designs, involving groups of patients affected by sensorimotor deficits, would be extremely valuable in delineating neuropsychological models of interpersonal coordination. As relevant groups, we suggest pathologies selectively involving the cerebellum and the basal ganglia, which have been proposed to serve dissociable functions in SMS (Cannon & Patel, 2021; Doya, 2000; Nozaradan et al., 2017; Schwartze et al., 2016). While these functions are often investigated by means of rhythmic cues, a dyadic setting may offer insights into the role of specific brain structures in interpersonal coordination. Persons affected by autistic spectrum disorder (ASD) represent another population of interest. Given the reported challenges in ASD with processing low-level information during interaction they offer a window into neurodivergence in social interaction (Ansuini, Podda, Battaglia, Veneselli, & Becchio, 2018; Montobbio et al., 2022; von der Lühe et al., 2016). Further investigation in this group may enhance our understanding of such dynamics in the neurotypical population across the lifespan.

  • *Pre-test / post-test designs*. In a test-retest longitudinal design, the drifting metronomes procedure can be utilized as a reference point to gauge modifications in interpersonal coordination following. One application could be to scrutinize changes in attractor dynamics following interventions aimed at improving dyadic cohesion and affiliation, such as a joint musical activity. Alternatively, this method could be used to investigate the plasticity of the attractor landscape in a dyadic system (Schöner et al., 1992): via either implicit or explicit learning, specific action patterns can be trained or reinforced such that they would consolidate as new attractors, reshaping the whole layout (Dhawale et al., 2017; Schöner et al., 1992). The effectiveness of such strategies could be tested through the following workflow: 1) initial scan of the entire attractor layout of a dyad, to establish a

[2] This principle was previously outlined in the Master's thesis of the author, titled "*Music as a framework, the musician as a model. Inferences on the bodily self*".

baseline profile of preferred patterns; 2) intervention in the dyad, or train a specific coordinative pattern (e.g., π/2); 3) repeat the scanning procedure following the intervention, to determine the extent of long-term plastic changes compared to the baseline.

Scaling to groups

In the ongoing development of this research, we are currently working on a substantial expansion of the dyadic paradigm to encompass the dynamics of larger groups, striving to incorporate further dimensions of social cognition into our investigation. We assert that the foundation of our paradigm, combined with the manipulation of diverse group configurations based on the social variables of experimental participants, can provide a controlled platform for the emergence of insightful *ingroup/outgroup dynamics*. Within a framework where the mutual attraction to synchronized behavior serves as an implicit behavioral indicator for cooperation (Marsh et al., 2009; Tzanaki, 2022), we ask whether and how coordination dynamics are related to facets of group affiliation, structure, relations, and identity[3].

Scaling to the group level brings an additional layer of complexity to our perspective and methodological approach. By focusing on the balance between cooperation and competition processes, we aim to explore how this operates with respect to ingroup-outgroup dynamics, as expressed in rhythmic configurations of coordinative patterns. Within this larger-scale approach, the field of ethnomusicology provides valuable background and a conceptual lens for interpreting these dynamics. For instance, cohesive adherence to a common rhythm can serve as a powerful means to affirm, separate, and preserve a group's identity (Lucas, Clayton, & Leante, 2011). Ethnographic research (Lucas, 2002b, 2002a, 2005) shows that for participants to music rituals, the preservation of boundaries between different groups often relies on maintaining distinct rhythmic patterns. Given the natural human inclination towards synchronization with others, resisting this drive becomes a powerful demonstration of a group's spiritual fortitude, expressed through its musical cohesion and competence at the rhythmic level. Therefore, when different communities come together in a ritual context, there is a compelling imperative to resist entrainment, further reinforcing group boundaries and identities. As

[3] This exploration forms the central theme of an upcoming project foreseen by the author of this dissertation.

we approach the commence of this group-level exploration, we anticipate that these insights will shed new light on the interplay of rhythmic patterns in social dynamics.

Computational modelling

Building on the substantial empirical groundwork presented throughout this dissertation, and possibly on its future expansion, we propose that the next leap forward involves further leveraging on computational modelling. Thus far, our work has primarily unfolded at a descriptive level, yielding behavioral observations under controlled conditions, from which we could infer the effect of variables of interest on the dynamics at play. By integrating this research with simulations of dyadic interactions, we can provide explanatory accounts for the phenomena we have reported. The main advantage of modelling is that it allows us to explicitly test the explanatory reach of the theoretical contenders we've outlined in this dissertation. With such an approach, we can generate and evaluate predictions starting from different stances and conclude which one better fits the empirical findings.

Transitioning our understanding from descriptive to predictive we can refine our existing data interpretations. Preliminary simulations, based on the setup of the drifting metronomes studies, are showing that leveraging the properties of the paradigm can provide unique insights into the mechanisms governing the system's dynamics. While this endeavor starts from a dyadic setup, it is meant to be easily scalable to any number of participants and multi-group configurations. This scalability goes in hand with the proposal of expanding our investigation to group dynamics, where simulations can inform and focus the experiments involving large number of participants. Our current efforts focus on extensions of the Kuramoto model, an ideal starting point given its relative simplicity and existing applications to human dyadic systems (Ole Adrian Heggli et al., 2019). We envision these improvements to function as a bridge between empirical observations and theoretical explanation. This will mark a decisive leap in improving our understanding of the mechanisms underlying interpersonal coordination.

Neurophysiology

In our explorations of neurophysiology, we have merely begun to scratch the surface, and the hyperscanning setup presents a wealth of analysis methods that remain largely untapped. Our argument has strongly advocated for a more interaction-centric approach

to modelling, coupling participants' effectors to oscillatory brain activity in a framework of cross-frequency coupling (M. Seeber et al., 2016; Martin Seeber, Scherer, Wagner, Solis-Escalante, & Müller-Putz, 2014; G. Zhou et al., 2016). Yet, there is merit in exploring the relationships between our methodology and more traditional hyperscanning analysis methods based on interbrain synchrony (Zamm et al., 2023).

Moreover, the non-dyadic experiments conducted so far have paved the way for the development of an analytical approach to neural entrainment. The ERFA protocol, as described in Chapter 7, is currently being tested in several studies on patients affected by multiple sclerosis and cerebellar lesions. This represents a first test for the replicability of our results, and a contribution to neuropsychological models of sensorimotor synchronization informed by neuroanatomical lesions. A next step for the application of this technique would consist of applying it to the hyperscanning setup, to investigate the dyadic interaction in terms of mutual frequency adjustment of neural components across participants. Adding another layer to our research, we propose to investigate the modulation of the proposed neural measures in relation to the attractor landscape[4]. Finally, our upcoming research on group dynamics will further expand on this, as it is prepared to scale the hyperscanning setup to incorporate small groups of participants (Astolfi et al., 2010).

[4] These suggestions are already set to be applied to an EEG hyperscanning dataset collected during the body-swap study detailed in Chapter 4.

Epilogue

The completion of this dissertation marks the culmination of nearly five years of dedicated research, a journey that crossed the fields of Arts Science and Psychology as part of a joint PhD program. The topic of rhythm is the cornerstone where the two domains converge.

The multi-dimensional nature of music, both as a form of joint action and as a cultural and emotional touchstone, makes it an extraordinary research domain for inferring more general principles of human interaction (P. E. Keller et al., 2014). Performance, dance, improvisation, playing in ensemble, are all social activities which require a delicate balance of precision and adaptability, making music a rich ecological context for studying rhythmic interpersonal coordination. This perspective delves into the psychological processes and neurophysiological mechanisms at play during a social interaction, exploring sensorimotor and cognitive processes, the representation and integration of self and other-related actions within and between individuals' brains, and the relationships between coordination and socio-cognitive constructs.

The current PhD dissertation moved from a complementary angle and tracked a different trajectory, originating in the most fundamental aspects of human interaction and streaming into the specific domain of systematic musicology (Marc Leman, 2007, 2016). It carves out a niche at the intersection of psychology and behavioral neuroscience, providing a multidisciplinary view on musical rhythm as a mediator for rich embodied interactions. In the spirit of art sciences, this dissertation embraces the bidirectionality of scientific discovery. Just as we learn about psychology and interaction from studying music, our understanding of human nature reciprocally enriches our comprehension and appreciation of music as a profound facet of human culture and sociality. This endeavor embodies the essence of a joint-PhD, reflecting the spirit of interdisciplinarity.

Dissecting complex human phenomena is not an easy task, and the path towards knowledge is a long, non-linear journey. Constantly oscillating back-and-forth between the analysis and the synthesis of its components, at times it feels like walking in circles. At times it feels out of balance. While the work of the past five years has been heavily analytical, these last few months have been a first attempt at synthesizing all my work together.

The pursuit of fundamental knowledge is a noble purpose. Yet, in the grand scheme of things, the true hope is that this work will find its place in the cross-pollination of scientific disciplines and practices, that neighboring fields can assimilate these insights, and that

applications will integrate them to make a meaningful impact. May fields such as music education, rehabilitation, artistic practices, and sports science all be informed by the bits of knowledge provided here, and by its future prospects.

With this wish for the future and the baggage of this experience, I am personally looking forward to starting a new cycle.

Appendix A

Supplementary materials

Chapter 2

All supplementary materials, including the audio file of the drifting metronomes, are downloadable on the website of *Scientific Reports*. Below, the full reference to the paper, accessible online in open access. Files can be found in the section '*Supplementary information'*.

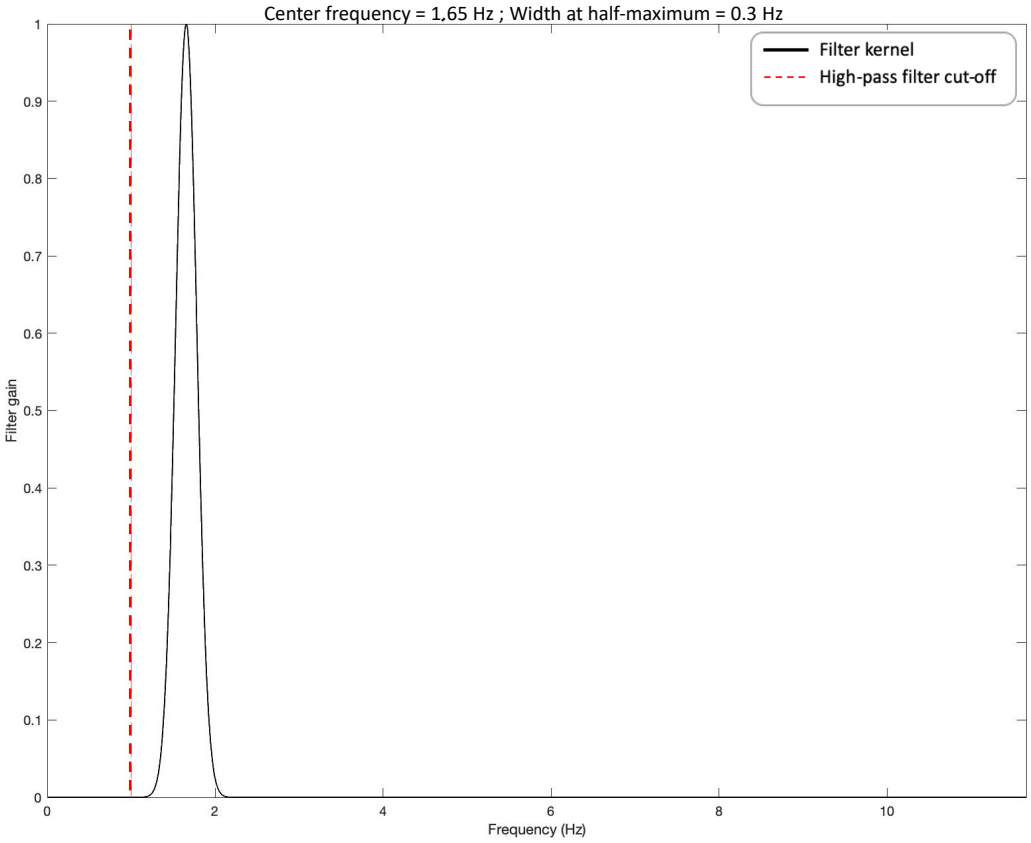
Rosso, M., Maes, P.J. & Leman, M. Modality-specific attractor dynamics in dyadic entrainment. Sci Rep 11, 18355 (2021). https://doi.org/10.1038/s41598-021-96054-8.

Chapter 5

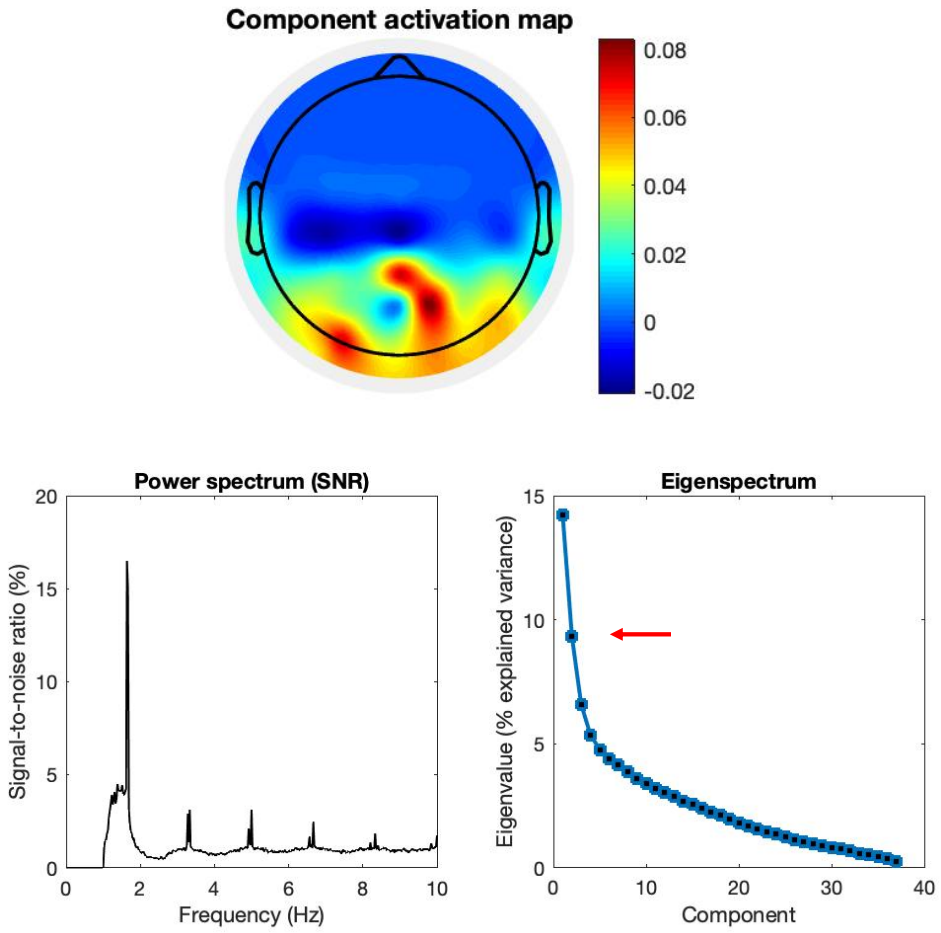
All supplementary materials are downloadable on the website of *Neuroimage*. Below, the full reference to the paper, accessible online in open access. Files can be found in the section '*Appendix. Supplementary materials.'*.

Rosso, M., Heggli, O. A., Maes, P. J., Vuust, P., & Leman, M. (2022). Mutual beta power modulation in dyadic entrainment. NeuroImage, 257, 119326.https://doi.org/10.1016/j.neuroimage.2022.119326

Chapter 6

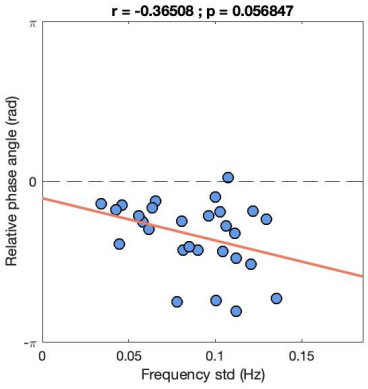


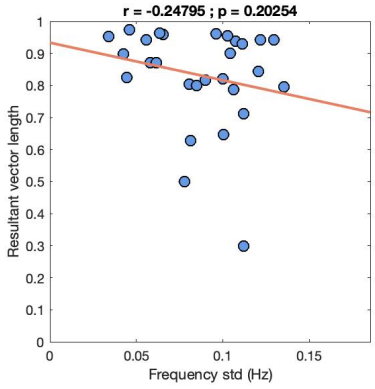
Supplementary Material 6. A Narrow-band filter kernel for GED. The figure shows the wavelet kernel in the frequency domain, centered on the target frequency of 1.65 Hz. In order to set the width of the Gaussian, we opted for a filter that would allow some extent of fluctuations around the centered frequency, without overlapping with the high-pass band filter (cut-off = 1 Hz). We found an optimal trade-off by setting the Gaussian width at half-maximum at 0.3 Hz. It is important to note that the correlations with the behavioral measures reported in the present work are invariant to such parameter. However, one should be aware that the scale of the stability index is inversely proportional to the width of the narrow-band filter. This is due to the fact that a wider filter allows for wider frequency fluctuations. Finally, the scale of the stability index is also affected by the filter shape. It was previously proposed that symmetric plateau-shaped filters should be preferred over a Gaussian when investigating frequency shifts, since the latter is biased towards the center frequency (Michael X. Cohen, 2014). However, we found that the correlations are invariant when comparing plateau-shaped FIR and Gaussian-shaped, as long as both are symmetrical. After verifying the invariance of the results, we opted for a Gaussian filter for the sake of parsimony and replicability: given that that the center frequency is constrained by the rate of the stimulation, the width is the only parameter to be defined by the data analyst.

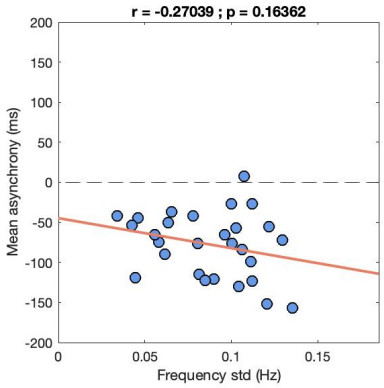


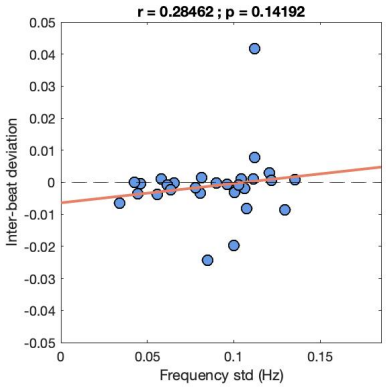
Supplementary Material 6. B. Component #2. As it can be evicted by the right-most plot, the second eigenvalue detaches to some extent in the eigenspectrum. We present here the activation map and the power spectrum (normalized to signal-to-noise ratio) of the component associated with the second eigenvalue. The spectral profile is clearly characterized by dominant peaks at entrained frequency and harmonics, which is to be expected given the spectral criterion adopted for the GED. It is noteworthy that the signal-to-noise ratio is considerably lower as compared to the first component. Critically, no meaningful pattern emerged from the topographical activation map.

Behavioral outcome measures ~ Stability index

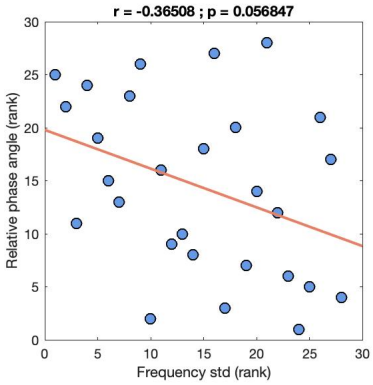


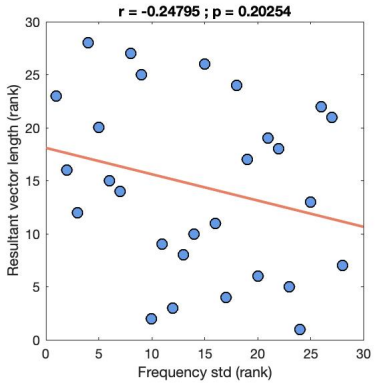


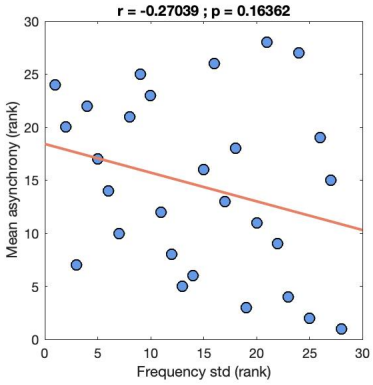


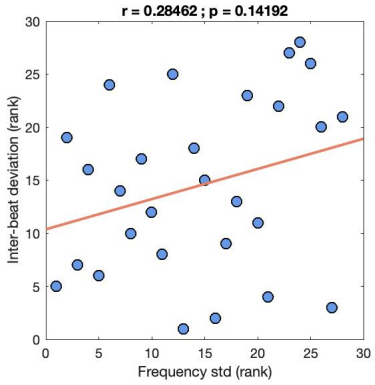


Behavioral outcome measures ~ Stability index
(Ranks)



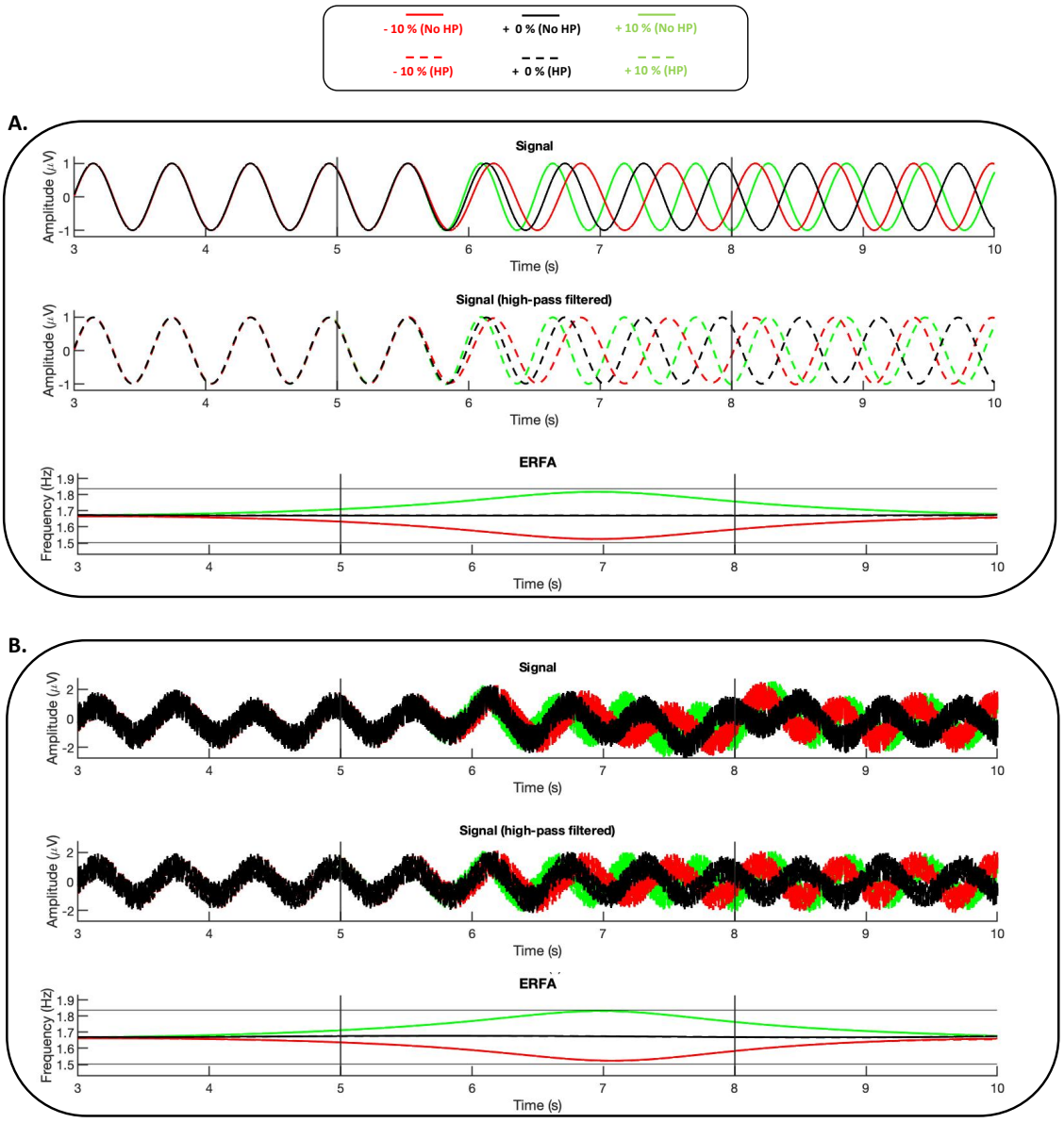




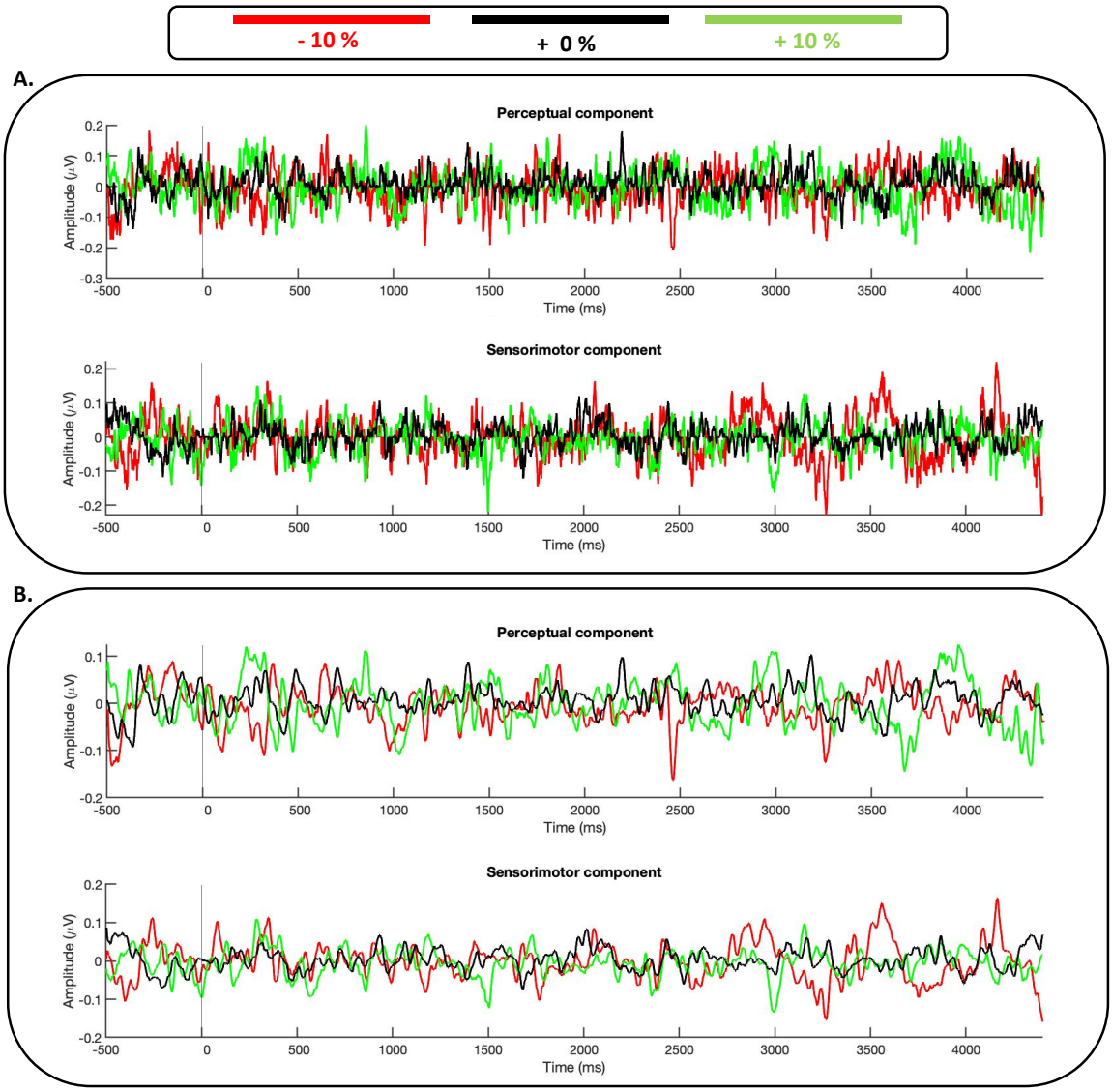


Supplementary Material 6. C. Component #2 (correlations with behavioral outcome measures). The stability index was computed from the second component, and its correlations with the behavioral outcome measures were tested. Correlations were considerably weaker than the ones for the component associated with the highest eigenvalue, as quantified by the Spearman correlation coefficients. Results are showed in the original scale and transformed to ranks. This evidence suggests that the first component alone is related to neural entrainment in the context of the experimental task. Acknowledging that we still cannot completely rule out the merging of different neural processes across components, the approach hereby proposed has proved to be valid in extracting one single component related to auditory-motor coupling.

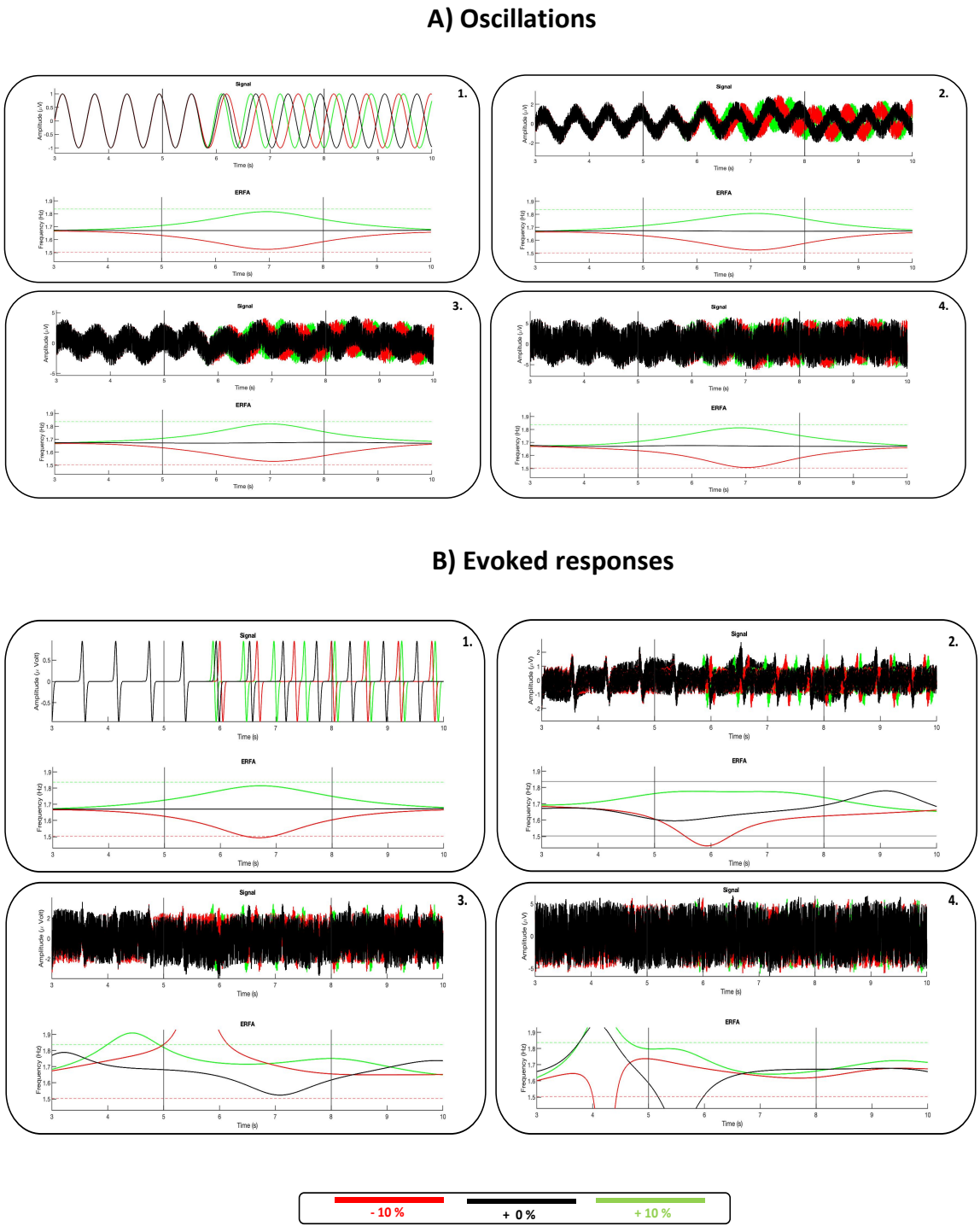
Chapter 7


- 10 % (No HP)
+ 0 % (No HP)
+ 10 % (No HP)
- 10 % (HP)
+ 0 % (HP)
+ 10 % (HP)
A.
Signal
Amplitude (μV)
Time (s)
Signal (high-pass filtered)
Amplitude (μV)
Time (s)
ERFA
Frequency (Hz)
Time (s)
B.
Signal
Amplitude (μV)
Time (s)
Signal (high-pass filtered)
Amplitude (μV)
Time (s)
ERFA
Frequency (Hz)
Time (s)

Supplementary Material 7. A. Simulated oscillations. *A) Simulation with and without high-pass filter.* The figure shows a comparison of a simulated oscillatory signal (1.667 Hz) entraining in three conditions of tempo change (+10%, 0%, -10%). The two upper subplots represent the signal activation, while the lowest subplot represents the ERFAs computed from the respective timeseries. In all subplots, continuous lines represent the signal and respective ERFAs without the high-pass filter, while dashed lines represent the same timeseries with the zero-phase high-pass filter (Butterworth, cutoff = 1 Hz; 6th order). In both cases, a narrow-band Gaussian filter (center = 1.67 Hz, full width at half-maximum = 0.3 Hz) was applied. The overlap between the dashed and continuous lines demonstrates that the 1 Hz high-pass filter does not influence the frequency activity at the target metronome frequency in the time windows of interest, leaving the ERFAs unbiased. Note that ERFAs from both scenarios are plotted together, but the dashed one is barely visible because of the overlap. *B) Simulation with and without high-pass filter, with noise.* The same timeseries are shown in the presence of noise (SNR white noise = 0.5; pink noise level = 0.3, exponential decay = 10). The overlap between the timeseries, even at lower signal-to-noise ratios, shows that the high-pass filter does not introduce any systematic bias in the analyzed time windows. The absence of distortion is further supported by the ERFA computed from our data during perturbation-free stimulation (black line in Figure 3 of the manuscript), which, being flat as expected at the reference frequency, serves as a valid baseline for assessing the effects of the perturbations.



Supplementary Material 7. B. Activation timeseries of perceptual and sensorimotor components. A) The figure shows the grand-average timeseries (N = 19) of the 'perceptual' and 'sensorimotor' components as extracted by GED. The ERFAs shown in Figure 3A and B in the main manuscript were computed from these timeseries. If evoked responses were prominent and passively followed the stimulation, they would be expected to track the tempo changes in the three levels of direction (+10%, 0%, -10%). However, no visible pattern of evoked responses corresponding to the tempo changes is observed, suggesting that the evoked responses are not a significant factor in the extracted components. It should be noted, however, that the task was not optimized for ERPs, finger-tapping was performed throughout its entire duration with no periods for baseline at rest, and the presented time window is particularly long for ERP standards. B) A sliding moving-average (30 ms window size) was applied to the same timeseries to facilitate the visibility of potential evoked response patterns. Even with smoothing, no evoked responses consistent with the stimulation pattern are observed.



Supplementary Material 7. C. Comparison of alternative models underlying ERFAs. The figures show two simulated scenarios based on Novembre and Iannetti (2018) tailored to our experimental design. The purpose of the simulation is to compare ERFAs as explained by alternative models for the underlying dynamics. In both scenarios, the upper subplot represents the simulated signal, while the lower subplot represents the ERFAs computed from the simulated timeseries. *A) Oscillations*: phase alignment, steady (+0%), speed-up (+10), and slow down (-10%) with respect to the original frequency. *B) Evoked responses*: passively following the corresponding changes in stimulation rate (+10%, +0%, -10%). For both models, we introduced varying levels of noise to assess the reliability of ERFAs with respect to the expected dynamics and the empirical evidence from experimental data.

1. No noise.
2. SNR (white noise) = 0.5; pink noise level = 0.3, exponential decay = 10.
3. SNR (white noise) = 0.2; pink noise level = 0.3, exponential decay = 10.
4. SNR (white noise) = 0.1; pink noise level = 0.3, exponential decay = 10.

In the absence of noise, both models A and B produce comparable expected ERFAs, suggesting that both models can explain the parabolic ERFA pattern under ideal conditions. As noise levels increase, however, model B (evoked responses) begins to degenerate and becomes dramatically distorted at lower SNRs, indicating that the model is less robust in less ideal scenarios which better approximates the reality of EEG signals. In contrast, the ERFAs generated by model A (oscillations) demonstrate remarkable robustness, maintaining consistency across all four noise levels. Furthermore, the ERFAs from model A closely resemble Figure 3B from the main manuscript (experimental data), providing evidence that an oscillatory model better explains the observed ERFAs in our study.

These findings support the conclusion that the oscillatory model (model A) is a more plausible explanation for the observed ERFAs in our data, as it demonstrates greater robustness to noise and closely aligns with the experimental findings presented in Figure 3B of the main manuscript.

Appendix B

List of acronyms

Concepts

HKB = Haken-Kelso-Buntz (model)

SMS = sensorimotor synchronization

Methods

ANOVA = analysis of variance

ART = aligned rank transform

DET = determinism

ERP = event-related potential

ERFA = event-related frequency adjustment

GED = generalized eigendecomposition

ICA = independent component analysis

JRQA = joint recurrence quantification analysis

PCA = principal component analysis

QoM = quantity of motion

RR = recurrence rate

RQA = recurrence quantification analysis

SEM = standard error of the mean

Tools

EEG = electroencephalography

EKG = electrocardiogram

EOG = electrooculogram

MIDI = musical instrument digital interface

MMSE = mini-mental state examination

VR = virtual reality

Institutions

ASIL = art science interaction lab

IPEM = Institute for Psychoacoustic and Electronic Music

PSITEC = Psychologie : Interactions, Temps, Emotions, Cognition

Appendix C

List of Figures

Appendix D: List of Tables

References

Afsar, O., Tirnakli, U., & Marwan, N. (2018). Recurrence Quantification Analysis at work: Quasi-periodicity based interpretation of gait force profiles for patients with Parkinson disease. *Scientific Reports*, *8*(1), 9102.

Agosta, F., Rocca, M. A., Pagani, E., Absinta, M., Magnani, G., Marcone, A., … Filippi, M. (2010). Sensorimotor network rewiring in mild cognitive impairment and Alzheimer's disease. *Human Brain Mapping*, *31*(4), 515–525.

Agres, K. R., Schaefer, R. S., Volk, A., van Hooren, S., Holzapfel, A., Dalla Bella, S., … Magee, W. L. (2021). Music, Computing, and Health: A Roadmap for the Current and Future Roles of Music Technology for Health Care and Well-Being. *Music & Science*, *4*, 2059204321997709.

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology: CB*, *14*(3), 257–262.

Allen, E. A., Liu, J., Kiehl, K. A., Gelernter, J., Pearlson, G. D., Perrone-Bizzozero, N. I., & Calhoun, V. D. (2011). Components of cross-frequency modulation in health and disease. *Frontiers in Systems Neuroscience*, *5*, 59.

Alperson, P. (1980). "musical time" and music as an "art of time." *Journal of Aesthetics and Art Criticism*, *38*(4), 407.

Andersen, L. M., & Dalal, S. S. (2021). The cerebellar clock: Predicting and timing somatosensory touch. *NeuroImage*, *238*, 118202.

Andersen, S. K., Fuchs, S., & Müller, M. M. (2011). Effects of feature-selective and spatial attention at different stages of visual processing. *Journal of Cognitive Neuroscience*, *23*(1), 238–246.

Andersen, S. K., Müller, M. M., & Martinovic, J. (2012). Bottom-up biases in feature-selective attention. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(47), 16953–16958.

Ansuini, C., Podda, J., Battaglia, F. M., Veneselli, E., & Becchio, C. (2018). One hand, two hands, two people: Prospective sensorimotor control in children with autism. *Developmental Cognitive Neuroscience*, *29*, 86–96.

Arnal, L. H., & Giraud, A.-L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, *16*(7), 390–398.

Aschersleben, G., & Prinz, W. (1995). Synchronizing actions with events: the role of sensory information. *Perception & Psychophysics*, *57*(3), 305–317.

Aschersleben, G., & Prinz, W. (1997). Delayed auditory feedback in synchronization. *Journal of Motor Behavior*, *29*(1), 35–46.

Aschersleben, Gisa. (2002). Temporal control of movements in sensorimotor synchronization. *Brain and Cognition*, *48*(1), 66–79.

Ashoori, A., Eagleman, D. M., & Jankovic, J. (2015). Effects of Auditory Rhythm and Music on Gait Disturbances in Parkinson's Disease. *Frontiers in Neurology*, *6*, 234.

Assaneo, M. F., Rimmele, J. M., Sanz Perl, Y., & Poeppel, D. (2021). Speaking rhythmically can shape hearing. *Nature Human Behaviour*, *5*(1), 71–82.

Assisi, C. G., Jirsa, V. K., & Kelso, J. A. S. (2005). Dynamics of multifrequency coordination using parametric driving: theory and experiment. *Biological Cybernetics*, *93*(1), 6–21.

Astolfi, L., Toppi, J., De Vico Fallani, F., Vecchiato, G., Salinari, S., Mattia, D., … Babiloni, F. (2010). Neuroelectrical hyperscanning measures simultaneous brain activity in humans. *Brain Topography*, *23*(3), 243–256.

Backwell, P., Jennions, M., Passmore, N., & Christy, J. (1998). Synchronized courtship in fiddler crabs. *Nature*, *391*(6662), 31–32.

Baird, A., & Samson, S. (2015). Music and dementia. *Progress in Brain Research*, *217*, 207–235.

Ballerini, M., Cabibbo, N., Candelier, R., Cavagna, A., Cisbani, E., Giardina, I., … Zdravkovic, V. (2008). Interaction ruling animal collective behavior depends on topological rather than metric distance: evidence from a field study. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(4), 1232–1237.

Bardy, B. G., Hoffmann, C. P., Moens, B., Leman, M., & Dalla Bella, S. (2015). Sound-induced stabilization of breathing and moving. *Annals of the New York Academy of Sciences*, *1337*(1), 94–100.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting Linear Mixed-Effects Models using lme4. Retrieved from http://arxiv.org/abs/1406.5823

Bavassi, L., Kamienkowski, J. E., Sigman, M., & Laje, R. (2017). Sensorimotor synchronization: neurophysiological markers of the asynchrony in a finger-tapping task. *Psychological Research*, *81*(1), 143–156.

Bavassi, M. L., Tagliazucchi, E., & Laje, R. (2013). Small perturbations in a finger-tapping task reveal inherent nonlinearities of the underlying error correction mechanism. *Human Movement Science*, *32*(1), 21–47.

Beek, P. J., Peper, C. E., & Daffertshofer, A. (2002). Modeling Rhythmic Interlimb Coordination: Beyond the Haken–Kelso–Bunz Model. *Brain and Cognition*, *48*(1), 149–165.

Bella, S. D., Benoit, C. E., Farrugia, N., Keller, P. E., & Obrig, H. (2017). Gait improvement via rhythmic stimulation in Parkinson's disease is linked to rhythmic skills. *Scientific Reports*. Retrieved from https://www.nature.com/articles/srep42005

Belz, M., Pyritz, L. W., & Boos, M. (2013). Spontaneous flocking in human groups. *Behavioural Processes*, *92*, 6–14.

Bengtsson, S. L., Ullén, F., Ehrsson, H. H., Hashimoto, T., Kito, T., Naito, E., … Sadato, N. (2009). Listening to rhythms activates motor and premotor cortices. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *45*(1), 62–71.

Berens, P. (2009). CircStat: A MATLAB Toolbox for Circular Statistics. *Journal of Statistical Software*, *31*, 1–21.

Bertelson, P., & Aschersleben, G. (2003). Temporal ventriloquism: crossmodal interaction on the time dimension: 1. Evidence from auditory–visual temporal order judgment. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *50*(1), 147–155.

Betti, V., Della Penna, S., de Pasquale, F., & Corbetta, M. (2021). Spontaneous Beta Band Rhythms in the Predictive Coding of Natural Stimuli. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *27*(2), 184–201.

Bevilacqua, F., Boyer, E. O., Françoise, J., Houix, O., Susini, P., Roby-Brami, A., & Hanneton, S. (2016). Sensori-Motor Learning with Movement Sonification: Perspectives from Recent Interdisciplinary Studies. *Frontiers in Neuroscience*, *10*, 385.

Bevilacqua, F., Peyre, I., Segalen, M., Pradat-Diehl, P., Marchand-Pauvert, V., & Roby-Brami, A. (2018, June 28). Exploring different movement sonification strategies for rehabilitation in clinical settings. *Proceedings of the 5th International Conference on Movement and Computing*, 1–6. Presented at the Genoa, Italy. New York, NY, USA: Association for Computing Machinery.

Biro, D., Sumpter, D. J. T., Meade, J., & Guilford, T. (2006). From compromise to leadership in pigeon homing. *Current Biology: CB*, *16*(21), 2123–2128.

Blakemore, S. J., Frith, C. D., & Wolpert, D. M. (1999). Spatio-temporal prediction modulates the perception of self-produced stimuli. *Journal of Cognitive Neuroscience*, *11*(5), 551–559.

Blakemore, Sarah Jayne, & Frith, C. (2003). Self-awareness and action. *Current Opinion in Neurobiology*, *13*(2), 219–224.

Blakemore, S.-J. (2017). Why can't you tickle yourself? In *The Anatomy of Laughter* (pp. 34–41). Routledge.

Boashash, B. (1992). Estimating and interpreting the instantaneous frequency of a signal. I. Fundamentals. *Proceedings of the IEEE*. Retrieved from https://ieeexplore.ieee.org/abstract/document/135376/

Botvinick, M., & Cohen, J. (1998). Rubber hands "feel" touch that eyes see. *Nature*, *391*(6669), 756.

Bourguignon, M., Jousmäki, V., Dalal, S. S., Jerbi, K., & De Tiège, X. (2019). Coupling between human brain activity and body movements: Insights from non-invasive electromagnetic recordings. *NeuroImage*, *203*, 116177.

Bourguignon, M., Piitulainen, H., De Tiège, X., Jousmäki, V., & Hari, R. (2015). Corticokinematic coherence mainly reflects movement-induced proprioceptive feedback. *NeuroImage*, *106*, 382–390.

Bradley, E., & Kantz, H. (2015). Nonlinear time-series analysis revisited. *Chaos* , *25*(9), 097610.

Bresin, R., Mancini, M., Elblaus, L., & Frid, E. (2020). Sonification of the self vs. sonification of the other: Differences in the sonification of performed vs. observed simple hand movements. *International Journal of Human-Computer Studies*, *144*, 102500.

Breska, A., & Deouell, L. Y. (2017). Neural mechanisms of rhythm-based temporal prediction: Delta phase-locking reflects temporal predictability but not rhythmic entrainment. *PLoS Biology*, *15*(2), e2001665.

Bron, R., & Furness, J. B. (2009). Rhythm of digestion: keeping time in the gastrointestinal tract. *Clinical and Experimental Pharmacology & Physiology*, *36*(10), 1041–1048.

Bruineberg, J., Kiverstein, J., & Rietveld, E. (2018). The anticipating brain is not a scientist: the free-energy principle from an ecological-enactive perspective. *Synthese*, *195*(6), 2417–2444.

Buhmann, J., Moens, B., Van Dyck, E., Dotov, D., & Leman, M. (2018). Optimizing beat synchronized running to music. *PloS One*, *13*(12), e0208702.

Burgess, A. P. (2013). On the interpretation of synchronization in EEG hyperscanning studies: a cautionary note. *Frontiers in Human Neuroscience*, *7*, 881.

Burr, D., Banks, M. S., & Morrone, M. C. (2009). Auditory dominance over vision in the perception of interval duration. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, *198*(1), 49–57.

Buzsáki, G., & Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science*, *304*(5679), 1926–1929.

Buzsáki, G., Logothetis, N., & Singer, W. (2013). Scaling brain size, keeping timing: evolutionary preservation of brain rhythms. *Neuron*, *80*(3), 751–764.

Cadena-Valencia, J., García-Garibay, O., Merchant, H., Jazayeri, M., & de Lafuente, V. (2018). Entrainment and maintenance of an internal metronome in supplementary motor area. *ELife*, *7*. doi:10.7554/eLife.38983

Cannon, J. J., & Patel, A. D. (2021). How Beat Perception Co-opts Motor Neurophysiology. *Trends in Cognitive Sciences*, *25*(2), 137–150.

Canolty, R. T., Edwards, E., Dalal, S. S., Soltani, M., Nagarajan, S. S., Kirsch, H. E., … Knight, R. T. (2006). High gamma power is phase-locked to theta oscillations in human neocortex. *Science*, *313*(5793), 1626–1628.

Chella, F., Marzetti, L., Stenroos, M., Parkkonen, L., Ilmoniemi, R. J., Romani, G. L., & Pizzella, V. (2019). The impact of improved MEG–MRI co-registration on MEG connectivity analysis. *NeuroImage*, *197*, 354–367.

Chemin, B., Mouraux, A., & Nozaradan, S. (2014). Body movement selectively shapes the neural representation of musical rhythms. *Psychological Science*, *25*(12), 2147–2159.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex* , *18*(12), 2844–2854.

Cheng, T.-H. Z., Creel, S. C., & Iversen, J. R. (2022). How Do You Feel the Rhythm: Dynamic Motor-Auditory Interactions Are Involved in the Imagination of Hierarchical Timing. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *42*(3), 500–512.

Cirelli, L. K. (2018). How interpersonal synchrony facilitates early prosocial behavior. *Current Opinion in Psychology*, *20*, 35–39.

Clark, A. (1999). Where brain, body, and world collide. *Cognitive Systems Research*, *1*(1), 5–17.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *The Behavioral and Brain Sciences*, *36*(3), 181–204.

Clark, A. (2017). Busting out: Predictive brains, embodied minds, and the puzzle of the evidentiary veil. *Nous* , *51*(4), 727–753.

Clayton, Jakubowski, & Eerola. (2020). Interpersonal entrainment in music performance: theory, method, and model. *Music Perception*. Retrieved from https://online.ucpress.edu/mp/article-abstract/38/2/136/114278

Clayton, M., Dueck, B., & Leante, L. (2013). *Experience and Meaning in Music Performance*. Oxford 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*, 1–82. Romanian Society for Ethnomusicology.

Cohen, E. E. A., Ejsmond-Frey, R., Knight, N., & Dunbar, R. I. M. (2010). Rowers' high: behavioural synchrony is correlated with elevated pain thresholds. *Biology Letters*, *6*(1), 106–108.

Cohen, M. X. (2014). *Analyzing neural time series data: theory and practice*. Retrieved from https://books.google.com/books?hl=en&lr=&id=rDKkAgAAQBAJ&oi=fnd&pg=PR5&dq=cohen+2014+analyzing+&ots=g9ditZ1VE_&sig=Hd_1AaGZvNMoKZMwi4J3FSezMvI

Cohen, Michael X. (2008). Assessing transient cross-frequency coupling in EEG data. *Journal of Neuroscience Methods*, *168*(2), 494–499.

Cohen, Michael X. (2014). Fluctuations in oscillation frequency control spike timing and coordinate neural networks. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(27), 8988–8998.

Cohen, Michael X. (2017). Where does EEG come from and what does it mean? *Trends in Neurosciences*, *40*(4), 208–218.

Cohen, Michael X. (2022). A tutorial on generalized eigendecomposition for denoising, contrast enhancement, and dimension reduction in multichannel electrophysiology. *NeuroImage*, *247*, 118809.

Cohen, Michael X., & Gulbinaite, R. (2017). Rhythmic entrainment source separation: Optimizing analyses of neural responses to rhythmic sensory stimulation. *NeuroImage*, *147*, 43–56.

Comstock, D. C., & Balasubramaniam, R. (2018). Neural responses to perturbations in visual and auditory metronomes during sensorimotor synchronization. *Neuropsychologia*, *117*, 55–66.

Comstock, D. C., Hove, M. J., & Balasubramaniam, R. (2018). Sensorimotor Synchronization With Auditory and Visual Modalities: Behavioral and Neural Differences. *Frontiers in Computational Neuroscience*, *12*, 53.

Couzin, I. D. (2009). Collective cognition in animal groups. *Trends in Cognitive Sciences*, *13*(1), 36–43.

Crapse, T. B., & Sommer, M. A. (2008). Corollary discharge across the animal kingdom. *Nature Reviews. Neuroscience*. Retrieved from https://www.nature.com/articles/nrn2457

Criscuolo, A., Schwartze, M., & Kotz, S. A. (2022). Cognition through the lens of a body–brain dynamic system. *Trends in Neurosciences*, *45*(9), 667–677.

Crombé, K., Denys, M., & Maes, P.-J. (2022). The role of a mechanical coupling in (spontaneous) interpersonal synchronization: A human version of Huygens' clock experiments. *Timing & Time Perception (Leiden, Netherlands)*, 1–20.

Cuijpers, L. S., Den Hartigh, R. J. R., Zaal, F. T. J. M., & de Poel, H. J. (2019). Rowing together: Interpersonal coordination dynamics with and without mechanical coupling. *Human Movement Science*, *64*, 38–46.

Cummins, F. (2009a). Rhythm as an affordance for the entrainment of movement. *Phonetica*, *66*(1–2), 15–28.

Cummins, F. (2009b). Rhythm as entrainment: The case of synchronous speech. *Journal of Phonetics*, *37*(1), 16–28.

Czeszumski, A., Eustergerling, S., Lang, A., Menrath, D., Gerstenberger, M., Schuberth, S., … König, P. (2020). Hyperscanning: A Valid Method to Study Neural Inter-brain Underpinnings of Social Interaction. *Frontiers in Human Neuroscience*, *14*, 39.

Dalla Bella, S., Farrugia, N., Benoit, C.-E., Begel, V., Verga, L., Harding, E., & Kotz, S. A. (2017). BAASTA: Battery for the Assessment of Auditory Sensorimotor and Timing Abilities. *Behavior Research Methods*, *49*(3), 1128–1145.

Davis, M. H., & Others. (1980). *A multidimensional approach to individual differences in empathy*. Retrieved from https://www.uv.es/friasnav/Davis_1980.pdf

de Cheveigné, A. (2000). A model of the perceptual asymmetry between peaks and troughs of frequency modulation. *The Journal of the Acoustical Society of America*, *107*(5 Pt 1), 2645–2656.

de Dreu, M. J., van der Wilk, A. S. D., Poppe, E., Kwakkel, G., & van Wegen, E. E. H. (2012). Rehabilitation, exercise therapy and music in patients with Parkinson's disease: a meta-analysis of the effects of music-based movement therapy on walking ability, balance and quality of life. *Parkinsonism & Related Disorders*, *18*, S114–S119.

de Guzman, G. C., & Tognoli, E. (2014). The human dynamic clamp as a paradigm for social interaction. *Proceedings of The*. Retrieved from https://www.pnas.org/content/111/35/E3726.short

De Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences*, *14*(10), 441–447.

de Pasquale, F., Corbetta, M., Betti, V., & Della Penna, S. (2018). Cortical cores in network dynamics. *NeuroImage*, *180*(Pt B), 370–382.

De Pasquale, F., Della Penna, S., Sporns, O., Romani, G. L., & Corbetta, M. (2016). A dynamic core network and global efficiency in the resting human brain. *Cerebral Cortex* , *26*(10), 4015–4033.

de Poel, H. J. (2016). Anisotropy and Antagonism in the Coupling of Two Oscillators: Concepts and Applications for Between-Person Coordination. *Frontiers in Psychology*, *7*, 1947.

Della Gatta, F., Garbarini, F., Puglisi, G., Leonetti, A., Berti, A., & Borroni, P. (2016). Decreased motor cortex excitability mirrors own hand disembodiment during the rubber hand illusion. *ELife*, *5*. doi:10.7554/eLife.14972

Dell'Anna, A., Fossataro, C., Burin, D., Bruno, V., Salatino, A., Garbarini, F., … Berti, A. (2018). Entrainment beyond embodiment. *Neuropsychologia*, *119*, 233–240.

Demany, L., & Clément, S. (1995). The perception of frequency peaks and troughs in wide frequency modulations. II. Effects of frequency register, stimulus uncertainty, and intensity. *The Journal of the Acoustical Society of America*, *97*(4), 2454–2459.

Demany, L., & Clément, S. (1997). The perception of frequency peaks and troughs in wide frequency modulations. IV. Effects of modulation waveform. *The Journal of the Acoustical Society of America*, *102*(5 Pt 1), 2935–2944.

Demany, L., & McAnally, K. I. (1994). The perception of frequency peaks and troughs in wide frequency modulations. *The Journal of the Acoustical Society of America*, *96*(2 Pt 1), 706–715.

Demos, A. P., & Chaffin, R. (2017). Removing obstacles to the analysis of movement in musical performance: Recurrence, mixed models, and surrogates. *The Routledge Companion to Embodied Music Interaction*, 341–349.

Demos, A. P., Chaffin, R., Begosh, K. T., Daniels, J. R., & Marsh, K. L. (2012). Rocking to the beat: effects of music and partner's movements on spontaneous interpersonal coordination. *Journal of Experimental Psychology. General*, *141*(1), 49–53.

Demos, A. P., Chaffin, R., & Kant, V. (2014). Toward a dynamical theory of body movement in musical performance. *Frontiers in Psychology*, *5*, 477.

Demos, A. P., Chaffin, R., & Logan, T. (2018). Musicians body sway embodies musical structure and expression: A recurrence-based approach. *Musicae Scientiae: The Journal of the European Society for the Cognitive Sciences of Music*, *22*(2), 244–263.

Demos, A. P., Layeghi, H., Wanderley, M. M., & Palmer, C. (2019). Staying together: A bidirectional delay-coupled approach to joint action. *Cognitive Science*, *43*(8), e12766.

Dermody, N., Wong, S., Ahmed, R., Piguet, O., Hodges, J. R., & Irish, M. (2016). Uncovering the Neural Bases of Cognitive and Affective Empathy Deficits in Alzheimer's Disease and the Behavioral-Variant of Frontotemporal Dementia. *Journal of Alzheimer's Disease: JAD*, *53*(3), 801–816.

Desmet, F., Lesaffre, M., Six, J., Ehrlé, N., & Samson, S. (2017). Multimodal analysis of synchronization data from patients with dementia. *25th Anniversary Conference of the European Society for the Cognitive Sciences of Music (ESCOM)*, 53–58. Ghent University.

Dhawale, A. K., Smith, M. A., & Ölveczky, B. P. (2017). The Role of Variability in Motor Learning. *Annual Review of Neuroscience*, *40*, 479–498.

Donoghue, T., Haller, M., Peterson, E. J., Varma, P., Sebastian, P., Gao, R., … Voytek, B. (2020). Parameterizing neural power spectra into periodic and aperiodic components. *Nature Neuroscience*, *23*(12), 1655–1665.

Dotov, D. G., de Cock, V. C., Geny, C., Driss, V., & Garrigue, G. (2017). Biologically-variable rhythmic auditory cues are superior to isochronous cues in fostering natural gait variability in Parkinson's disease. *Gait & Posture*. Retrieved from https://www.sciencedirect.com/science/article/pii/S0966636216305823

Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current Opinion in Neurobiology*, *10*(6), 732–739.

Dumas, G., Moreau, Q., Tognoli, E., & Kelso, J. A. S. (2020). The Human Dynamic Clamp Reveals the Fronto-Parietal Network Linking Real-Time Social Coordination and Cognition. *Cerebral Cortex* , *30*(5), 3271–3285.

Dumas, Guillaume, de Guzman, G. C., Tognoli, E., & Kelso, J. A. S. (2014). The human dynamic clamp as a paradigm for social interaction. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(35), E3726-34.

Dumas, Guillaume, Kelso, J. A. S., & Nadel, J. (2014). Tackling the social cognition paradox through multi-scale approaches. *Frontiers in Psychology*, *5*, 882.

Dunlap, K. (1910). Reaction to rhythmic stimuli with attempt to synchronize. *Psychological Review*, *17*(6), 399–416.

Eckmann, J.-P., Kamphorst, S. O., Ruelle, D., & Others. (1995). Recurrence plots of dynamical systems. *World Scientific Series on Nonlinear Science Series A*, *16*, 441–446.

Effenberg, A. O. (2005). Movement sonification: Effects on perception and action. *IEEE Multimedia*, *12*(2), 53–59.

Effenberg, Alfred O., Fehse, U., Schmitz, G., Krueger, B., & Mechling, H. (2016). Movement Sonification: Effects on Motor Learning beyond Rhythmic Adjustments. *Frontiers in Neuroscience*, *10*, 219.

Effenberg, Alfred O., & Schmitz, G. (2018). Acceleration and deceleration at constant speed: systematic modulation of motion perception by kinematic sonification. *Annals of the New York Academy of Sciences*, *1425*(1), 52–69.

Ehrsson, H. H. (2007). The experimental induction of out-of-body experiences. *Science*, *317*(5841), 1048.

Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of Vision*, *10*(1), 6.1-12.

Fait, S., Pighin, S., Passerini, A., Pavani, F., & Tentori, K. (2023). Sensory and multisensory reasoning: Is Bayesian updating modality-dependent? *Cognition*, *234*, 105355.

Farrer, C., & Frith, C. D. (2002). Experiencing Oneself vs Another Person as Being the Cause of an Action: The Neural Correlates of the Experience of Agency. *NeuroImage*, *15*(3), 596–603.

Feingold, J., Gibson, D. J., DePasquale, B., & Graybiel, A. M. (2015). Bursts of beta oscillation differentiate postperformance activity in the striatum and motor cortex of monkeys performing movement tasks. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(44), 13687–13692.

Ferreri, F., Vecchio, F., Vollero, L., Guerra, A., Petrichella, S., Ponzo, D., … Di Lazzaro, V. (2016). Sensorimotor cortex excitability and connectivity in Alzheimer's disease: A TMS-EEG Co-registration study. *Human Brain Mapping*, *37*(6), 2083–2096.

Fink, P. W., Foo, P., Jirsa, V. K., & Kelso, J. A. (2000). Local and global stabilization of coordination by sensory information. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, *134*(1), 9–20.

Fisher, N. I. (1995). *Statistical Analysis of Circular Data*. Cambridge University Press.

Fraser, A. M., & Swinney, H. L. (1986). Independent coordinates for strange attractors from mutual information. *Physical Review A: General Physics*, *33*(2), 1134–1140.

Freeman, W. J., Holmes, M. D., Burke, B. C., & Vanhatalo, S. (2003). Spatial spectra of scalp EEG and EMG from awake humans. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *114*(6), 1053–1068.

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, *9*(10), 474–480.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(1456), 815–836.

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews. Neuroscience*, *11*(2), 127–138.

Friston, K. (2011). Embodied inference: or "I think therefore I am, if I am what I think." In W. Tschacher (Ed.), *The implications of embodiment: Cognition and communication , (pp* (Vol. 279, pp. 89–125). Charlottesville, VA: Imprint Academic, x.

Friston, K. J., & Stephan, K. E. (2007). Free-energy and the brain. *Synthese*, *159*(3), 417–458.

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1521), 1211–1221.

Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology, Paris*, *100*(1–3), 70–87.

Friston, K., Mattout, J., & Kilner, J. (2011). Action understanding and active inference. *Biological Cybernetics*, *104*(1–2), 137–160.

Frith, C. D., & Frith, U. (1999). Interacting minds--a biological basis. *Science*, *286*(5445), 1692–1695.

Frith, Chris D. (2007). The social brain? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*(1480), 671–678.

Fujioka, T., Trainor, L., Large, E., & Ross, B. (2009). Beta and gamma rhythms in human auditory cortex during musical beat processing. *Annals of the New York Academy of Sciences*. Retrieved from

https://books.google.com/books?hl=en&lr=&id=w_sVnZjUsPoC&oi=fnd&pg=PA89&dq=fujioka+2009+beta&ots=z1gsjiskK6&sig=xfu5pF8XEs_MLhDKqWnIcBiR0uY

Fujioka, Takako, Ross, B., & Trainor, L. J. (2015). Beta-Band Oscillations Represent Auditory Beat and Its Metrical Hierarchy in Perception and Imagery. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(45), 15187–15198.

Fujioka, Takako, Trainor, L. J., Large, E. W., & Ross, B. (2012). Internalized Timing of Isochronous Sounds Is Represented in Neuromagnetic Beta Oscillations. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(5), 1791–1802.

Furuya, S., & Soechting, J. F. (2010). Role of auditory feedback in the control of successive keystrokes during piano playing. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, *204*(2), 223–237.

Galigani, M., Ronga, I., Fossataro, C., Bruno, V., Castellani, N., Rossi Sebastiano, A., … Garbarini, F. (2021). Like the back of my hand: Visual ERPs reveal a specific change detection mechanism for the bodily self. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *134*, 239–252.

Gallace, A., & Spence, C. (2006). Multisensory synesthetic interactions in the speeded classification of visual size. *Perception & Psychophysics*, *68*(7), 1191–1203.

Gallagher, S. (2001). The practice of mind. Theory, simulation or primary interaction? *Journal of Consciousness Studies*, *8*(5–6), 83–108.

Gallese, V. (2019). Embodied simulation. Its bearing on aesthetic experience and the dialogue between neuroscience and the humanities. *Gestalt Theory*, *41*(2), 113–127.

Gallese, V., & Sinigaglia, C. (2010). The bodily self as power for action. *Neuropsychologia*, *48*(3), 746–755.

Gallese, V., & Sinigaglia, C. (2011). What is so special about embodied simulation? *Trends in Cognitive Sciences*, *15*(11), 512–519.

Gámez, J., Mendoza, G., Prado, L., Betancourt, A., & Merchant, H. (2019). The amplitude in periodic neural state trajectories underlies the tempo of rhythmic tapping. *PLoS Biology*, *17*(4), e3000054.

Gan, L., Huang, Y., Zhou, L., Qian, C., & Wu, X. (2015). Synchronization to a bouncing ball with a realistic motion trajectory. *Scientific Reports*, *5*, 11974.

Ghai, S., Ghai, I., Schmitz, G., & Effenberg, A. O. (2018). Effect of rhythmic auditory cueing on parkinsonian gait: A systematic review and meta-analysis. *Scientific Reports*, *8*(1), 506.

Ghilain, M., Hobeika, L., Lesaffre, M., Schiaratura, L., Singh, A., Six, J., … Samson, S. (2020). Does a live performance impact synchronization to musical rhythm in cognitively impaired elderly? *Journal of Alzheimer's Disease: JAD*, *78*(3), 939–949.

Ghilain, M., Schiaratura, L., Singh, A., Lesaffre, M., & Samson, S. (2019). Is music special for people with dementia. *Music and Dementia: From Cognition to Therapy*, 24–40.

Gibson, J. J. (2014). *The Ecological Approach to Visual Perception: Classic Edition*. Psychology Press.

Goldman, W. P., Baty, J. D., Buckles, V. D., Sahrmann, S., & Morris, J. C. (1999). Motor dysfunction in mildly demented AD individuals without extrapyramidal signs. *Neurology*, *53*(5), 956–962.

Grabe, E., & Low, E. L. (2013). Durational variability in speech and the Rhythm Class Hypothesis. In *Laboratory Phonology 7*. Berlin, Boston: DE GRUYTER.

Grabli, D., Karachi, C., Welter, M.-L., Lau, B., Hirsch, E. C., Vidailhet, M., & François, C. (2012). Normal and pathological gait: what we learn from Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, *83*(10), 979–985.

Graf, R. (1994). *Self-rotation and spatial reference: The psychology of partner-centred localisations*. Frankfurt: Peter Lang.

Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, *19*(5), 893–906.

Grahn, J. A., Henry, M. J., & McAuley, J. D. (2011). FMRI investigation of cross-modal interactions in beat perception: audition primes vision, but not vice versa. *NeuroImage*, *54*(2), 1231–1243.

Greenfield, M. D. (1994a). COOPERATION AND CONFLICT IN THE EVOLUTION OF SIGNAL INTERACTIONS. *Annual Review of Ecology and Systematics*, *25*(1), 97–126.

Greenfield, M. D. (1994b). Synchronous and alternating choruses in insects and anurans: Common mechanisms and diverse functions. *American Zoologist*, *34*(6), 605–615.

GrÜnbaum, D. (1998). Schooling as a strategy for taxis in a noisy environment. *Evolutionary Ecology*, *12*(5), 503–522.

Guterstam, A., & Ehrsson, H. H. (2012). Disowning one's seen real body during an out-of-body illusion. *Consciousness and Cognition*, *21*(2), 1037–1042.

Haegens, S., & Zion Golumbic, E. (2018). Rhythmic facilitation of sensory processing: A critical review. *Neuroscience and Biobehavioral Reviews*, *86*, 150–165.

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, *5*(4), 382–385.

Haken, H., Kelso, J. A., & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics*, *51*(5), 347–356.

Hamilton, A. F. de C. (2021). Hyperscanning: Beyond the Hype. *Neuron*, *109*(3), 404–407.

Harrison, S. J., & Richardson, M. J. (2009). Horsing around: spontaneous four-legged coordination. *Journal of Motor Behavior*, *41*(6), 519–524.

Hasson, U., & Frith, C. D. (2016). Mirroring and beyond: coupled dynamics as a generalized framework for modelling social interactions. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *371*(1693). doi:10.1098/rstb.2015.0366

Hatsopoulos, N. G., & Suminski, A. J. (2011). Sensing with the motor cortex. *Neuron*, *72*(3), 477–487.

Haufe, S., Dähne, S., & Nikulin, V. V. (2014). Dimensionality reduction for the analysis of brain oscillations. *NeuroImage*, *101*, 583–597.

Heggli, Ole A., Konvalinka, I., Kringelbach, M. L., & Vuust, P. (2019). Musical interaction is influenced by underlying predictive models and musical expertise. *Scientific Reports*, *9*(1), 11048.

Heggli, Ole Adrian, Cabral, J., Konvalinka, I., Vuust, P., & Kringelbach, M. L. (2019). A Kuramoto model of self-other integration across interpersonal synchronization strategies. *PLoS Computational Biology*, *15*(10), e1007422.

Heggli, Ole Adrian, Konvalinka, I., Cabral, J., Brattico, E., Kringelbach, M. L., & Vuust, P. (2021). Transient brain networks underlying interpersonal strategies during synchronized action. *Social Cognitive and Affective Neuroscience*, *16*(1–2), 19–30.

Heggli, Ole Adrian, Konvalinka, I., Kringelbach, M. L., & Vuust, P. (2021). A metastable attractor model of self–other integration (MEAMSO) in rhythmic synchronization. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *376*(1835), 20200332.

Heilbron, M., & Chait, M. (2018). Great Expectations: Is there Evidence for Predictive Coding in Auditory Cortex? *Neuroscience*, *389*, 54–73.

Heyes, C., & Catmur, C. (2022). What Happened to Mirror Neurons? *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *17*(1), 153–168.

Hobeika, L., Ghilain, M., Schiaratura, L., Lesaffre, M., Huvent-Grelle, D., Puisieux, F., & Samson, S. (2021). Socio-emotional and motor engagement during musical activities in older adults with major neurocognitive impairment. *Scientific Reports*, *11*(1), 15291.

Hobeika, L., Ghilain, M., Schiaratura, L., Lesaffre, M., Puisieux, F., Huvent-Grelle, D., & Samson, S. (2022). The effect of the severity of neurocognitive disorders on emotional and motor responses to music. *Annals of the New York Academy of Sciences*, *1518*(1), 231–238.

Hohwy, J. (2013). *The Predictive Mind*. OUP Oxford.

Hove, M. J., Fairhurst, M. T., Kotz, S. A., & Keller, P. E. (2013). Synchronizing with auditory and visual rhythms: an fMRI assessment of modality differences and modality appropriateness. *NeuroImage*, *67*, 313–321.

Hove, M. J., Iversen, J. R., Zhang, A., & Repp, B. H. (2013). Synchronization with competing visual and auditory rhythms: bouncing ball meets metronome. *Psychological Research*, *77*(4), 388–398.

Hove, M. J., & Keller, P. E. (2010). Spatiotemporal relations and movement trajectories in visuomotor synchronization. *Music Perception*, *28*(1), 15–26.

Hove, M. J., & Risen, J. L. (2009). It's All in the Timing: Interpersonal Synchrony Increases Affiliation. *Social Cognition*, *27*(6), 949–960.

Hove, M. J., Spivey, M. J., & Krumhansl, C. L. (2010). Compatibility of motion facilitates visuomotor synchronization. *Journal of Experimental Psychology. Human Perception and Performance*, *36*(6), 1525–1534.

Huberth, M., Dauer, T., Nanou, C., Román, I., Gang, N., Reid, W., … Fujioka, T. (2019). Performance monitoring of self and other in a turn-taking piano duet: A dual-EEG study. *Social Neuroscience*, *14*(4), 449–461.

Huygens, C. (1888). Œuvres complètes de Christiaan Huygens publiées par la Société Hollandaise des Sciences, 22 voll. *The Hague: M. Nijhoff*.

Issartel, J., Marin, L., & Cadopi, M. (2007). Unintended interpersonal co-ordination: "can we march to the beat of our own drum?" *Neuroscience Letters*, *411*(3), 174–179.

Iversen, J. R., Patel, A. D., Nicodemus, B., & Emmorey, K. (2015). Synchronization to auditory and visual rhythms in hearing and deaf individuals. *Cognition*, *134*, 232–244.

Iversen, J. R., Repp, B. H., & Patel, A. D. (2009). Top-down control of rhythm perception modulates early auditory responses. *Annals of the New York Academy of Sciences*, *1169*(1), 58–73.

Ivry, R. B., & Keele, S. W. (1989). Timing functions of the cerebellum. *Journal of Cognitive Neuroscience*, *1*(2), 136–152.

Ivry, R. B., Keele, S. W., & Diener, H. C. (1988). Dissociation of the lateral and medial cerebellum in movement timing and movement execution. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, *73*(1), 167–180.

Ivry, Richard B., & Richardson, T. C. (2002). Temporal control and coordination: the multiple timer model. *Brain and Cognition*, *48*(1), 117–132.

Ivry, Richard B., & Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends in Cognitive Sciences*, *12*(7), 273–280.

Jacobsen, J.-H., Stelzer, J., Fritz, T. H., Chételat, G., La Joie, R., & Turner, R. (2015). Why musical memory can be preserved in advanced Alzheimer's disease. *Brain: A Journal of Neurology*, *138*(Pt 8), 2438–2450.

Jacoby, N., & McDermott, J. H. (2017). Integer Ratio Priors on Musical Rhythm Revealed Cross-culturally by Iterated Reproduction. *Current Biology: CB*, *27*(3), 359–370.

Jäncke, L., Loose, R., Lutz, K., Specht, K., & Shah, N. J. (2000). Cortical activations during paced finger-tapping applying visual and auditory pacing stimuli. *Brain Research. Cognitive Brain Research*, *10*(1–2), 51–66.

Jantzen, K. J., Ratcliff, B. R., & Jantzen, M. G. (2018). Cortical Networks for Correcting Errors in Sensorimotor Synchronization Depend on the Direction of Asynchrony. *Journal of Motor Behavior*, *50*(3), 235–248.

Johnson, W. S. (1899). RESEARCHES IN PRACTICE AND HABIT. *Science*, *10*(250), 527–529.

Kalckert, A., & Ehrsson, H. H. (2012). Moving a Rubber Hand that Feels Like Your Own: A Dissociation of Ownership and Agency. *Frontiers in Human Neuroscience*, *6*, 40.

Kalckert, A., & Ehrsson, H. H. (2014). The moving rubber hand illusion revisited: comparing movements and visuotactile stimulation to induce illusory ownership. *Consciousness and Cognition*, *26*, 117–132.

Kashiwase, Y., Matsumiya, K., Kuriki, I., & Shioiri, S. (2012). Time courses of attentional modulation in neural amplification and synchronization measured with steady-state visual-evoked potentials. *Journal of Cognitive Neuroscience*, *24*(8), 1779–1793.

Kato, M., & Konishi, Y. (2006). Auditory dominance in the error correction process: a synchronized tapping study. *Brain Research*, *1084*(1), 115–122.

Kay, M. M., Kay, M., & Wobbrock, J. O. (2020). Package "ARTool." *CRAN Repository (2016)*. doi:">10.1145/1978942.1978963>

Keehner, M., Guerin, S. A., Miller, M. B., Turk, D. J., & Hegarty, M. (2006). Modulation of neural activity by angle of rotation during imagined spatial transformations. *NeuroImage*, *33*(1), 391–398.

Keitel, C., Andersen, S. K., & Müller, M. M. (2010). Competitive effects on steady-state visual evoked potentials with frequencies in- and outside the α band. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, *205*(4), 489–495.

Keller, P. (2014). Ensemble performance: Interpersonal alignment of musical expression. In D. Fabian (Ed.), *Expressiveness in music performance: Empirical approaches across styles and cultures , (pp* (Vol. 383, pp. 260–282). New York, NY, US: Oxford University Press, xxx.

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 of London. Series B, Biological Sciences*, *369*(1658), 20130394.

Kelso, J. A. S. (1995). *Dynamic patterns: the self-organization of brain and behavior*. MIT Press.

Kelso, J. A. S., DelColle, J. D., & Schöner, G. (2018). Action-perception as a pattern formation process. *Attention and Performance*. doi:10.4324/9780203772010-5/action-perception-pattern-formation-process-kelso-delcolle-schã¶ner

Kelso, J. A. S., Scholz, J. P., & Schöner, G. (1986). Nonequilibrium phase transitions in coordinated biological motion: critical fluctuations. *Physics Letters. A*, *118*(6), 279–284.

Kelso, J. A. S., & Tognoli, E. (2009). Toward a complementary neuroscience: metastable coordination dynamics of the brain. *Downward Causation and the Neurobiology of Free*. Retrieved from https://link.springer.com/chapter/10.1007/978-3-642-03205-9_6

Kelso, J. A. Scott. (2009). Synergies: atoms of brain and behavior. *Advances in Experimental Medicine and Biology*, *629*, 83–91.

Kelso, J. A. Scott, Dumas, G., & Tognoli, E. (2013). Outline of a general theory of behavior and brain coordination. *Neural Networks: The Official Journal of the International Neural Network Society*, *37*, 120–131.

Kepecs, A., & Fishell, G. (2014). Interneuron cell types are fit to function. *Nature*, *505*(7483), 318–326.

Kessler, K. (2000). Spatial cognition and verbal localisations: A connectionist model for the interpretation of spatial prepositions. *Wiesbaden: Deutscher Universitäts-Verlag*.

Kessler, Klaus, & Thomson, L. A. (2010). The embodied nature of spatial perspective taking: embodied transformation versus sensorimotor interference. *Cognition*, *114*(1), 72–88.

Kilner, J. M., Paulignan, Y., & Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Current Biology: CB*, *13*(6), 522–525.

Kilner, James M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cognitive Processing*, *8*(3), 159–166.

Kilteni, K., Groten, R., & Slater, M. (2012). The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual*. Retrieved from https://direct.mit.edu/pvar/article-abstract/21/4/373/18838

Kilteni, Konstantina, Maselli, A., Kording, K. P., & Slater, M. (2015). Over my fake body: body ownership illusions for studying the multisensory basis of own-body perception. *Frontiers in Human Neuroscience*, *9*, 141.

Kimura, K., Ogata, T., & Miyake, Y. (2020). Effects of a partner's tap intervals on an individual's timing control increase in slow-tempo dyad synchronisation using finger-tapping. *Scientific Reports*, *10*(1), 8237.

Kingsbury, L., Huang, S., Wang, J., Gu, K., Golshani, P., Wu, Y. E., & Hong, W. (2019). Correlated Neural Activity and Encoding of Behavior across Brains of Socially Interacting Animals. *Cell*, *178*(2), 429-446.e16.

Kleinfeld, D., Ahissar, E., & Diamond, M. E. (2006). Active sensation: insights from the rodent vibrissa sensorimotor system. *Current Opinion in Neurobiology*, *16*(4), 435–444.

Kliger Amrani, A., & Zion Golumbic, E. (2022). Memory-Paced Tapping to Auditory Rhythms: Effects of Rate, Speech, and Motor Engagement. *Journal of Speech, Language, and Hearing Research: JSLHR*, *65*(3), 923–939.

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: theory and data. *Psychology of Learning and Motivation*, *54*, 59–101.

Knoblich, G., & Sebanz, N. (2008). Evolving intentions for social interaction: from entrainment to joint action. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *363*(1499), 2021–2031.

Koban, L., Ramamoorthy, A., & Konvalinka, I. (2019). Why do we fall into sync with others? Interpersonal synchronization and the brain's optimization principle. *Social Neuroscience*, *14*(1), 1–9.

Konvalinka, I., Vuust, P., Roepstorff, A., & Frith, C. D. (2010). Follow you, follow me: continuous mutual prediction and adaptation in joint tapping. *Quarterly Journal of Experimental Psychology* , *63*(11), 2220–2230.

Koshimori, Y., & Thaut, M. H. (2018). Future perspectives on neural mechanisms underlying rhythm and music based neurorehabilitation in Parkinson's disease. *Ageing Research Reviews*, *47*, 133–139.

Kugler, P. N., & Turvey, M. T. (2015). *Information, Natural Law, and the Self-Assembly of Rhythmic Movement*. Routledge.

Lagarde, J. (2013). Challenges for the understanding of the dynamics of social coordination. *Frontiers in Neurorobotics*, *7*, 18.

Lakatos, P., Barczak, A., Neymotin, S. A., McGinnis, T., Ross, D., Javitt, D. C., & O'Connell, M. N. (2016). Global dynamics of selective attention and its lapses in primary auditory cortex. *Nature Neuroscience*, *19*(12), 1707–1717.

Lakatos, P., Chen, C.-M., O'Connell, M. N., Mills, A., & Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron*, *53*(2), 279–292.

Lakatos, P., Gross, J., & Thut, G. (2019). A New Unifying Account of the Roles of Neuronal Entrainment. *Current Biology: CB*, *29*(18), R890–R905.

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science*, *320*(5872), 110–113.

Lakatos, P., Musacchia, G., O'Connel, M. N., Falchier, A. Y., Javitt, D. C., & Schroeder, C. E. (2013). The spectrotemporal filter mechanism of auditory selective attention. *Neuron*, *77*(4), 750–761.

Lakatos, P., O'Connell, M. N., Barczak, A., Mills, A., Javitt, D. C., & Schroeder, C. E. (2009). The leading sense: supramodal control of neurophysiological context by attention. *Neuron*, *64*(3), 419–430.

Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., & Schroeder, C. E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *Journal of Neurophysiology*, *94*(3), 1904–1911.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, *106*(1), 119–159.

Large, E. W., & Snyder, J. S. (2009). Pulse and meter as neural resonance. *Annals of the New York Academy of Sciences*, *1169*, 46–57.

Latash, M. L. (2010). Motor synergies and the equilibrium-point hypothesis. *Motor Control*, *14*(3), 294–322.

Lazzerini Ospri, L., Prusky, G., & Hattar, S. (2017). Mood, the Circadian System, and Melanopsin Retinal Ganglion Cells. *Annual Review of Neuroscience*, *40*, 539–556.

Lee, J. H., Whittington, M. A., & Kopell, N. J. (2013). Top-down beta rhythms support selective attention via interlaminar interaction: a model. *PLoS Computational Biology*, *9*(8), e1003164.

Legrain, V., Iannetti, G. D., Plaghki, L., & Mouraux, A. (2011). The pain matrix reloaded: a salience detection system for the body. *Progress in Neurobiology*, *93*(1), 111–124.

Leman, M., & Godøy, R. I. (2010). Why study musical gestures? *Musical Gestures*. doi:10.4324/9780203863411-7/study-musical-gestures-marc-leman-rolf-inge-godã¸y

Leman, Marc. (2007). *Embodied Music Cognition and Mediation Technology*. MIT Press.

Leman, Marc. (2016). *The Expressive Moment: How Interaction (with Music) Shapes Human Empowerment*. MIT Press.

Lenc, T., Keller, P. E., Varlet, M., & Nozaradan, S. (2018). Neural tracking of the musical beat is enhanced by low-frequency sounds. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(32), 8221–8226.

Lenc, T., Merchant, H., Keller, P. E., Honing, H., Varlet, M., & Nozaradan, S. (2021). Mapping between sound, brain and behaviour: four-level framework for understanding rhythm processing in humans and non-human primates. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *376*(1835), 20200325.

Leow, L.-A., Waclawik, K., & Grahn, J. A. (2018). The role of attention and intention in synchronization to music: effects on gait. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, *236*(1), 99–115.

Lesaffre, M., Moens, B., & Desmet, F. (2017). Monitoring music and movement interaction in people with dementia. *The Routledge Companion To*. doi:10.4324/9781315621364-33/monitoring-music-movement-interaction-people-dementia-micheline-lesaffre-bart-moens-frank-desmet

Lesaffre, Micheline. (2018). Investigating Embodied Music Cognition for Health and Well-Being. In R. Bader (Ed.), *Springer Handbook of Systematic Musicology* (pp. 779–791). Berlin, Heidelberg: Springer Berlin Heidelberg.

Liu, D., Liu, S., Liu, X., Zhang, C., Li, A., Jin, C., … Zhang, X. (2018). Interactive Brain Activity: Review and Progress on EEG-Based Hyperscanning in Social Interactions. *Frontiers in Psychology*, *9*, 1862.

London, J. (2012). *Hearing in time: Psychological aspects of musical meter*. Retrieved from https://books.google.ca/books?hl=en&lr=&id=8vUJCAAAQBAJ&oi=fnd&pg=PP1&ots=HHq6vPeFPB&sig=eOT5D3bxJKCSpLmEfaE_PA_Nfkw

López, S. L., & Laje, R. (2019). Spatiotemporal perturbations in paced finger tapping suggest a common mechanism for the processing of time errors. *Scientific Reports*, *9*(1), 17814.

Lorås, H., Aune, T. K., Ingvaldsen, R., & Pedersen, A. V. (2019). Interpersonal and intrapersonal entrainment of self-paced tapping rate. *PloS One*, *14*(7), e0220505.

Lorenzoni, V., Maes, P.-J., Van den Berghe, P., De Clercq, D., de Bie, T., & Leman, M. (2018, June 28). A biofeedback music-sonification system for gait retraining. *Proceedings of the 5th International Conference on Movement and Computing*, 1–5. Presented at the Genoa, Italy. New York, NY, USA: Association for Computing Machinery.

Lorenzoni, V., Staley, J., Marchant, T., Onderdijk, K. E., Maes, P.-J., & Leman, M. (2019). The sonic instructor: A music-based biofeedback system for improving weightlifting technique. *PloS One*, *14*(8), e0220915.

Lorenzoni, V., Van den Berghe, P., Maes, P.-J., De Bie, T., De Clercq, D., & Leman, M. (2019). Design and validation of an auditory biofeedback system for modification of running parameters. *Journal on Multimodal User Interfaces*, *13*(3), 167–180.

Lucas, G. (2002a). Musical Rituals of Afro-Brazilian Religious Groups within the Ceremonies of Congado. *Yearbook for Traditional Music*, *34*, 115–128.

Lucas, G. (2002b). *Os sons do Rosário: o congado mineiro dos Arturos e Jatobá* (Vol. 86). Editoria UFMG.

Lucas, G. (2005). *Música e tempo nos rituais do congado mineiro dos Arturos e do Jatobá*.

Lucas, G., Clayton, M., & Leante, L. (2011). Inter-group entrainment in Afro-Brazilian Congado ritual. *Empirical Musicology Review: EMR*, *6*(2), 75–102.

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique, second edition*. MIT Press.

Luo, L., & Flanagan, J. G. (2007). Development of continuous and discrete neural maps. *Neuron*, *56*(2), 284–300.

MacDougall, H. G., & Moore, S. T. (2005). Marching to the beat of the same drummer: the spontaneous tempo of human locomotion. *Journal of Applied Physiology*, *99*(3), 1164–1173.

Maeda, F., Kanai, R., & Shimojo, S. (2004). Changing pitch induced visual motion illusion. *Current Biology: CB*, *14*(23), R990-1.

Maes, P.-J., Buhmann, J., & Leman, M. (2016). 3Mo: A Model for Music-Based Biofeedback. *Frontiers in Neuroscience*, *10*, 548.

Maes, P.-J., Giacofci, M., & Leman, M. (2015). Auditory and motor contributions to the timing of melodies under cognitive load. *Journal of Experimental Psychology. Human Perception and Performance*, *41*(5), 1336–1352.

Maes, P.-J., Lorenzoni, V., & Six, J. (2019). The SoundBike: musical sonification strategies to enhance cyclists' spontaneous synchronization to external music. *Journal on Multimodal User Interfaces*, *13*(3), 155–166.

Maister, L., Slater, M., Sanchez-Vives, M. V., & Tsakiris, M. (2015). Changing bodies changes minds: owning another body affects social cognition. *Trends in Cognitive Sciences*, *19*(1), 6–12.

Marks, L. E. (2004). Cross-Modal Interactions in Speeded Classification. In G. A. Calvert (Ed.), *The handbook of multisensory processes , (pp* (Vol. 915, pp. 85–105). Cambridge, MA, US: Boston Review, xvii.

Marsh, K. L., Isenhower, R. W., Richardson, M. J., Helt, M., Verbalis, A. D., Schmidt, R. C., & Fein, D. (2013). Autism and social disconnection in interpersonal rocking. *Frontiers in Integrative Neuroscience*, *7*, 4.

Marsh, K. L., Richardson, M. J., Baron, R. M., & Schmidt, R. C. (2006). Contrasting approaches to perceiving and acting with others. *Ecological Psychology: A Publication of the International Society for Ecological Psychology*, *18*(1), 1–38.

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.

Marwan, N., Carmen Romano, M., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports*, *438*(5), 237–329.

Marwan, N., & Kurths, J. (2002). Nonlinear analysis of bivariate data with cross recurrence plots. *Physics Letters. A*, *302*(5), 299–307.

Marwan, N., Wessel, N., Meyerfeldt, U., Schirdewan, A., & Kurths, J. (2002). Recurrence-plot-based measures of complexity and their application to heart-rate-variability data. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, *66*(2 Pt 2), 026702.

Maselli, A., & Slater, M. (2013). The building blocks of the full body ownership illusion. *Frontiers in Human Neuroscience*, *7*, 83.

Matthews, T. E., Witek, M. A. G., Heggli, O. A., Penhune, V. B., & Vuust, P. (2019). The sensation of groove is affected by the interaction of rhythmic and harmonic complexity. *PloS One*, *14*(1), e0204539.

Matthews, T. E., Witek, M. A. G., Lund, T., Vuust, P., & Penhune, V. B. (2020). The sensation of groove engages motor and reward networks. *NeuroImage*, *214*, 116768.

May, M. (2004). Imaginal perspective switches in remembered environments: transformation versus interference accounts. *Cognitive Psychology*, *48*(2), 163–206.

Mazzoni, P., & Krakauer, J. W. (2006). An implicit plan overrides an explicit strategy during visuomotor adaptation. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *26*(14), 3642–3645.

McAuley, J. D. (2010). Tempo and Rhythm. In M. Riess Jones, R. R. Fay, & A. N. Popper (Eds.), *Music Perception* (pp. 165–199). New York, NY: Springer New York.

McPherson, T., Berger, D., Alagapan, S., & Fröhlich, F. (2018). Intrinsic Rhythmicity Predicts Synchronization-Continuation Entrainment Performance. *Scientific Reports*, *8*(1), 11782.

Menaker, W., & Menaker, A. (1959). Lunar periodicity in human reproduction: a likely unit of biological time. *American Journal of Obstetrics and Gynecology*, *77*(4), 905–914.

Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M., & Fitch, W. T. (2015). Finding the beat: a neural perspective across humans and non-human primates. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*(1664), 20140093.

Merker, B. H., Madison, G. S., & Eckerdal, P. (2009). On the role and origin of isochrony in human rhythmic entrainment. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *45*(1), 4–17.

Middlebrooks, J. C., & Green, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology*, *42*, 135–159.

Miller, K. J., Leuthardt, E. C., Schalk, G., Rao, R. P. N., Anderson, N. R., Moran, D. W., … Ojemann, J. G. (2007). Spectral changes in cortical surface potentials during motor movement. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(9), 2424–2432.

Mirman, D. (2017). *Growth Curve Analysis and Visualization Using R*. CRC Press.

Mitchell, J. P., Banaji, M. R., & Macrae, C. N. (2005). The link between social cognition and self-referential thought in the medial prefrontal cortex. *Journal of Cognitive Neuroscience*, *17*(8), 1306–1315.

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.

Miyata, K., Varlet, M., Miura, A., Kudo, K., & Keller, P. E. (2018). Interpersonal visual interaction induces local and global stabilisation of rhythmic coordination. *Neuroscience Letters*, *682*, 132–136.

Miyata, K., Varlet, M., Miura, A., Kudo, K., & Keller, P. E. (2021). Vocal interaction during rhythmic joint action stabilizes interpersonal coordination and individual movement timing. *Journal of Experimental Psychology. General*, *150*(2), 385–394.

Moelants, D. (2002). PREFERRED TEMPO RECONSIDERED. Retrieved June 4, 2023, from https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=0762d815903c643b587250cfc1f4659c9da9c4a5

Moens, B. (2018). D-jogger: An interactive music system for gait synchronisation with applications for sports and rehabilitation. Retrieved May 31, 2023, from https://biblio.ugent.be/publication/8551818/file/8551819

Moens, Bart, Muller, C., van Noorden, L., Franěk, M., Celie, B., Boone, J., … Leman, M. (2014). Encouraging spontaneous synchronisation with D-Jogger, an adaptive music player that aligns movement and music. *PloS One*, *9*(12), e114234.

Montobbio, N., Cavallo, A., Albergo, D., Ansuini, C., Battaglia, F., Podda, J., … Becchio, C. (2022). Intersecting kinematic encoding and readout of intention in autism. *Proceedings of the National Academy of Sciences*, *119*(5), e2114648119.

Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: examining temporal ventriloquism. *Brain Research. Cognitive Brain Research*, *17*(1), 154–163.

Morillon, B., Arnal, L. H., Schroeder, C. E., & Keitel, A. (2019). Prominence of delta oscillatory rhythms in the motor cortex and their relevance for auditory and speech perception. *Neuroscience and Biobehavioral Reviews*, *107*, 136–142.

Morillon, B., & Baillet, S. (2017). Motor origin of temporal predictions in auditory attention. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(42), E8913–E8921.

Morillon, B., Hackett, T. A., Kajikawa, Y., & Schroeder, C. E. (2015). Predictive motor control of sensory dynamics in auditory active sensing. *Current Opinion in Neurobiology*, *31*, 230–238.

Morillon, B., Schroeder, C. E., & Wyart, V. (2014). Motor contributions to the temporal precision of auditory attention. *Nature Communications*, *5*, 5255.

Moumddjian, L. (2020). Auditory-motor coupling to musical and non-musical rhythms during walking in persons with Multiple Sclerosis. Retrieved May 30, 2023, from https://biblio.ugent.be/publication/8659179/file/8659180.pdf

Moumdjian, L., Buhmann, J., Willems, I., Feys, P., & Leman, M. (2018). Entrainment and Synchronization to Auditory Stimuli During Walking in Healthy and Neurological Populations: A Methodological Systematic Review. *Frontiers in Human Neuroscience*, *12*, 263.

Moumdjian, L., Maes, P.-J., Dalla Bella, S., Decker, L. M., Moens, B., Feys, P., & Leman, M. (2020). Detrended fluctuation analysis of gait dynamics when entraining to music and metronomes at different tempi in persons with multiple sclerosis. *Scientific Reports*, *10*(1), 12934.

Moumdjian, L., Moens, B., Maes, P.-J., Van Geel, F., Ilsbroukx, S., Borgers, S., … Feys, P. (2019). Continuous 12 min walking to music, metronomes and in silence: Auditory-motor coupling and its effects on perceived fatigue, motivation and gait in persons with multiple sclerosis. *Multiple Sclerosis and Related Disorders*, *35*, 92–99.

Moumdjian, L., Moens, B., Maes, P.-J., Van Nieuwenhoven, J., Van Wijmeersch, B., Leman, M., & Feys, P. (2019). Walking to Music and Metronome at Various Tempi in Persons With Multiple Sclerosis: A Basis for Rehabilitation. *Neurorehabilitation and Neural Repair*, *33*(6), 464–475.

Moumdjian, L., Moens, B., Vanzeir, E., De Klerck, B., Feys, P., & Leman, M. (2019). A model of different cognitive processes during spontaneous and intentional coupling to music in multiple sclerosis. *Annals of the New York Academy of Sciences*, *1445*(1), 27–38.

Moumdjian, L., Rosso, M., Moens, B., De Weerdt, N., Leman, M., & Feys, P. (2022). A case-study of a person with multiple sclerosis and cerebellar ataxia synchronizing finger-taps and foot-steps to music and metronomes. *Neuroimmunology Reports*, *2*, 100101.

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

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

Nessler, J. A., & Gilliland, S. J. (2010). Kinematic analysis of side-by-side stepping with intentional and unintentional synchronization. *Gait & Posture*, *31*(4), 527–529.

Neuper, C., & Pfurtscheller, G. (2001). Event-related dynamics of cortical rhythms: frequency-specific features and functional correlates. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *43*(1), 41–58.

Niarchou, M., Gustavson, D. E., Sathirapongsasuti, J. F., Anglada-Tort, M., Eising, E., Bell, E., … Gordon, R. L. (2022). Genome-wide association study of musical beat synchronization demonstrates high polygenicity. *Nature Human Behaviour*, *6*(9), 1292–1309.

Nijhuis, P., Keller, P. E., Nozaradan, S., & Varlet, M. (2021). Dynamic modulation of cortico-muscular coupling during real and imagined sensorimotor synchronisation. *NeuroImage*, *238*, 118209.

Nobre, A. C., & van Ede, F. (2018). Anticipated moments: temporal structure in attention. *Nature Reviews. Neuroscience*, *19*(1), 34–48.

Nombela, C., Rae, C. L., Grahn, J. A., Barker, R. A., Owen, A. M., & Rowe, J. B. (2013). How often does music and rhythm improve patients' perception of motor symptoms in Parkinson's disease? *Journal of Neurology*, *260*(5), 1404–1405.

Norcia, A. M., Appelbaum, L. G., Ales, J. M., Cottereau, B. R., & Rossion, B. (2015). The steady-state visual evoked potential in vision research: A review. *Journal of Vision*, *15*(6), 4.

Novembre, G., & Iannetti, G. D. (2018). [Review of *Tagging the musical beat: Neural entrainment or event-related potentials?*]. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(47), E11002–E11003. National Acad Sciences.

Novembre, G., & Iannetti, G. D. (2021). Hyperscanning Alone Cannot Prove Causality. Multibrain Stimulation Can. *Trends in Cognitive Sciences*, *25*(2), 96–99.

Novembre, G., Knoblich, G., Dunne, L., & Keller, P. E. (2017). Interpersonal synchrony enhanced through 20 Hz phase-coupled dual brain stimulation. *Social Cognitive and Affective Neuroscience*. doi:10.1093/scan/nsw172

Novembre, G., Mitsopoulos, Z., & Keller, P. E. (2019). Empathic perspective taking promotes interpersonal coordination through music. *Scientific Reports*, *9*(1), 12255.

Novembre, G., Sammler, D., & Keller, P. E. (2016). Neural alpha oscillations index the balance between self-other integration and segregation in real-time joint action. *Neuropsychologia*, *89*, 414–425.

Novembre, G., Ticini, L. F., Schütz-Bosbach, S., & Keller, P. E. (2012). Distinguishing self and other in joint action. Evidence from a musical paradigm. *Cerebral Cortex* , *22*(12), 2894–2903.

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 of the United States of America*, *108*(52), 20947–20952.

Noy, L., Levit-Binun, N., & Golland, Y. (2015). Being in the zone: physiological markers of togetherness in joint improvisation. *Frontiers in Human Neuroscience*, *9*, 187.

Nozaradan, S., Peretz, I., & Keller, P. E. (2016). Individual Differences in Rhythmic Cortical Entrainment Correlate with Predictive Behavior in Sensorimotor Synchronization. *Scientific Reports*, *6*, 20612.

Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(28), 10234–10240.

Nozaradan, S., Peretz, I., & Mouraux, A. (2012). Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(49), 17572–17581.

Nozaradan, S., Schwartze, M., Obermeier, C., & Kotz, S. A. (2017). Specific contributions of basal ganglia and cerebellum to the neural tracking of rhythm. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *95*, 156–168.

Nozaradan, S., Zerouali, Y., Peretz, I., & Mouraux, A. (2015). Capturing with EEG the neural entrainment and coupling underlying sensorimotor synchronization to the beat. *Cerebral Cortex* , *25*(3), 736–747.

Obleser, J., Henry, M. J., & Lakatos, P. (2017). What do we talk about when we talk about rhythm? *PLoS Biology*, *15*(9), e2002794.

Obleser, J., & Kayser, C. (2019). Neural Entrainment and Attentional Selection in the Listening Brain. *Trends in Cognitive Sciences*, *23*(11), 913–926.

Oh, H., Braun, A. R., Reggia, J. A., & Gentili, R. J. (2019). Fronto-parietal mirror neuron system modeling: Visuospatial transformations support imitation learning independently of imitator perspective. *Human Movement Science*, *65*. doi:10.1016/j.humov.2018.05.013

Okano, M., Shinya, M., & Kudo, K. (2017). Paired Synchronous Rhythmic Finger Tapping without an External Timing Cue Shows Greater Speed Increases Relative to Those for Solo Tapping. *Scientific Reports*, *7*, 43987.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113.

Onderdijk, K. E. (2022). Come together: Exploring unity in music interaction through agency and social connectedness. Retrieved May 31, 2023, from https://biblio.ugent.be/publication/8763490/file/8763491

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*, 156869.

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

Paladino, M.-P., Mazzurega, M., Pavani, F., & Schubert, T. W. (2010). Synchronous multisensory stimulation blurs self-other boundaries. *Psychological Science*, *21*(9), 1202–1207.

Palisson, J., Roussel-Baclet, C., Maillet, D., Belin, C., Ankri, J., & Narme, P. (2015). Music enhances verbal episodic memory in Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology*, *37*(5), 503–517.

Palmer, C., & Demos, A. P. (2021). *Are we in time? How predictive coding and dynamical systems explain musical synchrony*. doi:10.31219/osf.io/8enxy

Palmer, C., & Demos, A. P. (2022). Are We in Time? How Predictive Coding and Dynamical Systems Explain Musical Synchrony. *Current Directions in Psychological Science*, *31*(2), 147–153.

Parise, C. V., & Spence, C. (2009). "When birds of a feather flock together": synesthetic correspondences modulate audiovisual integration in non-synesthetes. *PloS One*, *4*(5), e5664.

Park, H., Ince, R. A. A., Schyns, P. G., Thut, G., & Gross, J. (2015). Frontal top-down signals increase coupling of auditory low-frequency oscillations to continuous speech in human listeners. *Current Biology: CB*, *25*(12), 1649–1653.

Parseihian, G., Gondre, C., Aramaki, M., Ystad, S., & Kronland-Martinet, R. (2016). Comparison and Evaluation of Sonification Strategies for Guidance Tasks. *IEEE Transactions on Multimedia*, *18*(4), 674–686.

Pavani, F., Spence, C., & Driver, J. (2000). Visual capture of touch: out-of-the-body experiences with rubber gloves. *Psychological Science*, *11*(5), 353–359.

Pecenka, N., & Keller, P. E. (2011). The role of temporal prediction abilities in interpersonal sensorimotor synchronization. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, *211*(3–4), 505–515.

Peck, T. C., Doan, M., Bourne, K. A., & Good, J. J. (2018). The Effect of Gender Body-Swap Illusions on Working Memory and Stereotype Threat. *IEEE Transactions on Visualization and Computer Graphics*, *24*(4), 1604–1612.

Peck, T. C., Good, J. J., & Bourne, K. A. (2020). Inducing and Mitigating Stereotype Threat Through Gendered Virtual Body-Swap Illusions. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–13). New York, NY, USA: Association for Computing Machinery.

Peck, T. C., Seinfeld, S., Aglioti, S. M., & Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and Cognition*, *22*(3), 779–787.

Petkova, V. I., Björnsdotter, M., Gentile, G., Jonsson, T., Li, T.-Q., & Ehrsson, H. H. (2011). From part- to whole-body ownership in the multisensory brain. *Current Biology: CB*, *21*(13), 1118–1122.

Petkova, V. I., & Ehrsson, H. H. (2008). If I were you: perceptual illusion of body swapping. *PloS One*, *3*(12), e3832.

Petkova, V. I., Khoshnevis, M., & Ehrsson, H. H. (2011). The perspective matters! Multisensory integration in ego-centric reference frames determines full-body ownership. *Frontiers in Psychology*, *2*, 35.

Pfurtscheller, G. (1981). Central beta rhythm during sensorimotor activities in man. *Electroencephalography and Clinical Neurophysiology*, *51*(3), 253–264.

Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *110*(11), 1842–1857.

Phillips-Silver, J., Aktipis, C. A., & Bryant, G. A. (2010). The ecology of entrainment: Foundations of coordinated rhythmic movement. *Music Perception*, *28*(1), 3–14.

Phillips-Silver, J., & Keller, P. E. (2012). Searching for roots of entrainment and joint action in early musical interactions. *Frontiers in Human Neuroscience*, *6*, 26.

Piitulainen, H., Bourguignon, M., De Tiège, X., Hari, R., & Jousmäki, V. (2013). Corticokinematic coherence during active and passive finger movements. *Neuroscience*, *238*, 361–370.

Pikovsky, A., Rosenblum, M., Kurths, J., & Synchronization, A. (2001). A universal concept in nonlinear sciences. *Self*, *2*, 3.

Polak, R., Jacoby, N., & Fischinger, T. (2018). Rhythmic prototypes across cultures: A comparative study of tapping synchronization. *Music Perception: An*. Retrieved from https://online.ucpress.edu/mp/article-abstract/36/1/1/92062

Praamstra, P., Turgeon, M., Hesse, C. W., Wing, A. M., & Perryer, L. (2003). Neurophysiological correlates of error correction in sensorimotor-synchronization. *NeuroImage*, *20*(2), 1283–1297.

Prinz, W. (1990). A Common Coding Approach to Perception and Action. In O. Neumann & W. Prinz (Eds.), *Relationships Between Perception and Action: Current Approaches* (pp. 167–201). Berlin, Heidelberg: Springer Berlin Heidelberg.

Prinz, Wolfgang. (2013). Common coding. In *Encyclopedia of the mind* (pp. 161–163). New Delhi, India: SAGE.

Radeau, M., & Bertelson, P. (1974). The after-effects of ventriloquism. *The Quarterly Journal of Experimental Psychology*, *26*(1), 63–71.

Raible, F., Takekata, H., & Tessmar-Raible, K. (2017). An Overview of Monthly Rhythms and Clocks. *Frontiers in Neurology*, *8*, 189.

Rajendran, V. G., & Schnupp, J. W. H. (2019). [Review of *Frequency tagging cannot measure neural tracking of beat or meter*]. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(8), 2779–2780. National Acad Sciences.

Ramus, F., Nespor, M., & Mehler, J. (2000). Correlates of linguistic rhythm in the speech signal. *Cognition*, *75*(1), AD3–AD30.

Repp, B. H. (2001a). Phase correction, phase resetting, and phase shifts after subliminal timing perturbations in sensorimotor synchronization. *Journal of Experimental Psychology. Human Perception and Performance*, *27*(3), 600–621.

Repp, B. H. (2001b). Processes underlying adaptation to tempo changes in sensorimotor synchronization. *Human Movement Science*, *20*(3), 277–312.

Repp, Bruno H. (2000). Compensation for subliminal timing perturbations in perceptual-motor synchronization. *Psychological Research*, *63*(2), 106–128.

Repp, Bruno H. (2005). Sensorimotor synchronization: a review of the tapping literature. *Psychonomic Bulletin & Review*, *12*(6), 969–992.

Repp, Bruno H., & Keller, P. E. (2004). Adaptation to tempo changes in sensorimotor synchronization: effects of intention, attention, and awareness. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *57*(3), 499–521.

Repp, Bruno H., & Penel, A. (2002). Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences. *Journal of Experimental Psychology. Human Perception and Performance*, *28*(5), 1085–1099.

Repp, Bruno H., & Penel, A. (2004). Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychological Research*, *68*(4), 252–270.

Repp, Bruno H., & Su, Y.-H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin & Review*, *20*(3), 403–452.

Richardson, Dale, & Shockley. (2008). Synchrony and swing in conversation: Coordination, temporal dynamics, and communication. *Embodied Communication In*. Retrieved from https://books.google.com/books?hl=en&lr=&id=txVREAAAQBAJ&oi=fnd&pg=PA75&dq=Richardson,+Dale,+%26+Shockley,+2008&ots=d1CViHE_aR&sig=WlyCyJCOiEDwLJJMogcgaP6MHik

Richardson, M. J., Marsh, K. L., & Baron, R. M. (2007). Judging and actualizing intrapersonal and interpersonal affordances. *Journal of Experimental Psychology. Human Perception and Performance*, *33*(4), 845–859.

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

Richardson, M. J., Marsh, K. L., & Schmidt, R. C. (2005). Effects of visual and verbal interaction on unintentional interpersonal coordination. *Journal of Experimental Psychology. Human Perception and Performance*, *31*(1), 62–79.

Riley, M. A., Richardson, M. J., Shockley, K., & Ramenzoni, V. C. (2011). Interpersonal synergies. *Frontiers in Psychology*, *2*, 38.

Rimmele, J. M., Morillon, B., Poeppel, D., & Arnal, L. H. (2018). Proactive Sensing of Periodic and Aperiodic Auditory Patterns. *Trends in Cognitive Sciences*, *22*(10), 870–882.

Rivera, A. M., & Huberman, A. D. (2020). [Review of *Neuroscience: A Chromatic Retinal Circuit Encodes Sunrise and Sunset for the Brain*]. *Current biology: CB*, *30*(7), R316–R318. Elsevier.

Rodger, M. W. M., & Craig, C. M. (2011). Timing movements to interval durations specified by discrete or continuous sounds. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, *214*(3), 393–402.

Roman, I. R., Washburn, A., Large, E. W., Chafe, C., & Fujioka, T. (2019). Delayed feedback embedded in perception-action coordination cycles results in anticipation behavior

during synchronized rhythmic action: A dynamical systems approach. *PLoS Computational Biology*, *15*(10), e1007371.

Rosenblum, M., Pikovsky, A., Kurths, J., Schäfer, C., & Tass, P. A. (2001). Chapter 9 Phase synchronization: From theory to data analysis. In F. Moss & S. Gielen (Eds.), *Handbook of Biological Physics* (Vol. 4, pp. 279–321). North-Holland.

Rossion, B., Prieto, E. A., Boremanse, A., Kuefner, D., & Van Belle, G. (2012). A steady-state visual evoked potential approach to individual face perception: effect of inversion, contrast-reversal and temporal dynamics. *NeuroImage*, *63*(3), 1585–1600.

Rosso, M., Heggli, O. A., Maes, P. J., Vuust, P., & Leman, M. (2022). Mutual beta power modulation in dyadic entrainment. *NeuroImage*, *257*, 119326.

Rosso, M., Leman, M., & Moumdjian, L. (2021). Neural Entrainment Meets Behavior: The Stability Index as a Neural Outcome Measure of Auditory-Motor Coupling. *Frontiers in Human Neuroscience*, *15*, 668918.

Rosso, M., Maes, P. J., & Leman, M. (2021). Modality-specific attractor dynamics in dyadic entrainment. *Scientific Reports*, *11*(1), 18355.

Rulkov, N. F., Sushchik, M. M., Tsimring, L. S., & Abarbanel, H. D. (1995). Generalized synchronization of chaos in directionally coupled chaotic systems. *Physical Review. E, Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, *51*(2), 980–994.

Rutledge, D. N., & Jouan-Rimbaud Bouveresse, D. (2013). Independent Components Analysis with the JADE algorithm. *Trends in Analytical Chemistry: TRAC*, *50*, 22–32.

Sakamoto, M., Ando, H., & Tsutou, A. (2013). Comparing the effects of different individualized music interventions for elderly individuals with severe dementia. *International Psychogeriatrics / IPA*, *25*(5), 775–784.

Saleh, M., Reimer, J., Penn, R., Ojakangas, C. L., & Hatsopoulos, N. G. (2010). Fast and slow oscillations in human primary motor cortex predict oncoming behaviorally relevant cues. *Neuron*, *65*(4), 461–471.

Salenius, S., Schnitzler, A., Salmelin, R., Jousmäki, V., & Hari, R. (1997). Modulation of human cortical rolandic rhythms during natural sensorimotor tasks. *NeuroImage*, *5*(3), 221–228.

Salmelin, R., & Hari, R. (1994). Spatiotemporal characteristics of sensorimotor neuromagnetic rhythms related to thumb movement. *Neuroscience*, *60*(2), 537–550.

Salminen, N. H., Tiitinen, H., & May, P. J. C. (2012). Auditory spatial processing in the human cortex. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *18*(6), 602–612.

Samaha, J., & Cohen, M. X. (2022). Power spectrum slope confounds estimation of instantaneous oscillatory frequency. *NeuroImage*, *250*, 118929.

Samson, S., Clément, S., Narme, P., Schiaratura, L., & Ehrlé, N. (2015). Efficacy of musical interventions in dementia: methodological requirements of nonpharmacological trials. *Annals of the New York Academy of Sciences*, *1337*, 249–255.

Savage, P. E., Loui, P., Tarr, B., Schachner, A., Glowacki, L., Mithen, S., & Fitch, W. T. (2020). Music as a coevolved system for social bonding. *The Behavioral and Brain Sciences*, *44*, e59.

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

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

Schmidt, R. C., & Turvey, M. T. (1994). Phase-entrainment dynamics of visually coupled rhythmic movements. *Biological Cybernetics*, *70*(4), 369–376.

Schmidt, Richard C., & Richardson, M. J. (2008). Dynamics of Interpersonal Coordination. In A. Fuchs & V. K. Jirsa (Eds.), *Coordination: Neural, Behavioral and Social Dynamics* (pp. 281–308). Berlin, Heidelberg: Springer Berlin Heidelberg.

Schöner, G. (1989). Learning and recall in a dynamic theory of coordination patterns. *Biological Cybernetics*, *62*(1), 39–54.

Schöner, G., & Kelso, J. A. (1988). A synergetic theory of environmentally-specified and learned patterns of movement coordination. I. Relative phase dynamics. *Biological Cybernetics*, *58*(2), 71–80.

Schöner, G., Zanone, P. G., & Kelso, J. A. (1992). Learning as change of coordination dynamics: theory and experiment. *Journal of Motor Behavior*, *24*(1), 29–48.

Schönherr, J., & Westra, E. (2019). Beyond 'interaction': How to understand social effects on social cognition. *The British Journal for the Philosophy of Science*, *70*(1), 27–52.

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9–18.

Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., & Lakatos, P. (2010). Dynamics of Active Sensing and perceptual selection. *Current Opinion in Neurobiology*, *20*(2), 172–176.

Schwartze, M., Keller, P. E., & Kotz, S. A. (2016). Spontaneous, synchronized, and corrective timing behavior in cerebellar lesion patients. *Behavioural Brain Research*, *312*, 285–293.

Scott Kelso, J. A., Engstrom, D. A., & Engstrom, D. (2006). *The Complementary Nature*. MIT Press.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences*, *10*(2), 70–76.

Sebanz, N., & Knoblich, G. (2009). Prediction in joint action: what, when, and where. *Topics in Cognitive Science*, *1*(2), 353–367.

Seeber, M., Scherer, R., & Müller-Putz, G. R. (2016). EEG oscillations are modulated in different behavior-related networks during rhythmic finger movements. *Journal of Neuroscience*. Retrieved from https://www.jneurosci.org/content/36/46/11671?utm_source=TrendMD&utm_medium=cpc&utm_campaign=JNeurosci_TrendMD_1

Seeber, Martin, Scherer, R., Wagner, J., Solis-Escalante, T., & Müller-Putz, G. R. (2014). EEG beta suppression and low gamma modulation are different elements of human upright walking. *Frontiers in Human Neuroscience*, *8*, 485.

Seibold, P. (2021). *Sine fitting (https://www.mathworks.com/matlabcentral/fileexchange/66793-sine-fitting), MATLAB Central File Exchange*. Retrieved from https://www.mathworks.com/matlabcentral/fileexchange/66793-sine-fitting)

Serino, A., Alsmith, A., Costantini, M., Mandrigin, A., Tajadura-Jimenez, A., & Lopez, C. (2013). Bodily ownership and self-location: components of bodily self-consciousness. *Consciousness and Cognition*, *22*(4), 1239–1252.

Shadmehr, R., Smith, M. A., & Krakauer, J. W. (2010). Error correction, sensory prediction, and adaptation in motor control. *Annual Review of Neuroscience*, *33*, 89–108.

Shaffer, F., McCraty, R., & Zerr, C. L. (2014). A healthy heart is not a metronome: an integrative review of the heart's anatomy and heart rate variability. *Frontiers in Psychology*, *5*, 1040.

Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions. What you see is what you hear. *Nature*, *408*(6814), 788.

Sherman, M. A., Lee, S., Law, R., Haegens, S., Thorn, C. A., Hämäläinen, M. S., … Jones, S. R. (2016). Neural mechanisms of transient neocortical beta rhythms: Converging evidence from humans, computational modeling, monkeys, and mice. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(33), E4885-94.

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.

Shuler, M. G. H. (2016). Timing in the visual cortex and its investigation. *Current Opinion in Behavioral Sciences*, *8*, 73–77.

Sihvonen, A. J., Särkämö, T., Leo, V., Tervaniemi, M., Altenmüller, E., & Soinila, S. (2017). Music-based interventions in neurological rehabilitation. *Lancet Neurology*, *16*(8), 648–660.

Slater, M., Perez-Marcos, D., Ehrsson, H. H., & Sanchez-Vives, M. V. (2008). Towards a digital body: the virtual arm illusion. *Frontiers in Human Neuroscience*, *2*, 6.

Slater, M., Perez-Marcos, D., Ehrsson, H. H., & Sanchez-Vives, M. V. (2009). Inducing illusory ownership of a virtual body. *Frontiers in Neuroscience*, *3*(2), 214–220.

Smith, M. A., Ghazizadeh, A., & Shadmehr, R. (2006). Interacting adaptive processes with different timescales underlie short-term motor learning. *PLoS Biology*, *4*(6), e179.

Spence, C. (2011). Crossmodal correspondences: a tutorial review. *Attention, Perception & Psychophysics*, *73*(4), 971–995.

Stepp, N., & Turvey, M. T. (2010). On Strong Anticipation. *Cognitive Systems Research*, *11*(2), 148–164.

Stergiou, N., & Decker, L. M. (2011). Human movement variability, nonlinear dynamics, and pathology: is there a connection? *Human Movement Science*, *30*(5), 869–888.

Strogatz, S., Friedman, M., Mallinckrodt, A. J., & McKay, S. (1994). Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering. *Computer Physics Communications*, *8*(5), 532.

Strogatz, S. H. (2000). From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators. *Physica D. Nonlinear Phenomena*, *143*(1), 1–20.

Stupacher, J., Hove, M. J., Novembre, G., Schütz-Bosbach, S., & Keller, P. E. (2013). Musical groove modulates motor cortex excitability: a TMS investigation. *Brain and Cognition*, *82*(2), 127–136.

Styns, F., van Noorden, L., Moelants, D., & Leman, M. (2007). Walking on music. *Human Movement Science*, *26*(5), 769–785.

Tajadura-Jiménez, A., Grehl, S., & Tsakiris, M. (2012). The other in me: interpersonal multisensory stimulation changes the mental representation of the self. *PloS One*, *7*(7), e40682.

Takens, F. (1981). Detecting strange attractors in turbulence. *Dynamical Systems and Turbulence, Warwick 1980*, 366–381. Springer Berlin Heidelberg.

Teki, S., Grube, M., Kumar, S., & Griffiths, T. D. (2011). Distinct neural substrates of duration-based and beat-based auditory timing. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(10), 3805–3812.

Thériault, R., Olson, J. A., Krol, S. A., & Raz, A. (2021). Body swapping with a Black person boosts empathy: Using virtual reality to embody another. *Quarterly Journal of Experimental Psychology* , *74*(12), 2057–2074.

Thomson, M., Murphy, K., & Lukeman, R. (2018). Groups clapping in unison undergo size-dependent error-induced frequency increase. *Scientific Reports*, *8*(1), 808.

Tognoli, E., & Kelso, J. A. S. (2014). The metastable brain. *Neuron*, *81*(1), 35–48.

Tognoli, E., & Kelso, J. A. S. (2015). The coordination dynamics of social neuromarkers. *Frontiers in Human Neuroscience*, *9*, 563.

Tognoli, E., Lagarde, J., DeGuzman, G. C., & Kelso, J. A. S. (2007). The phi complex as a neuromarker of human social coordination. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(19), 8190–8195.

Tognoli, E., Zhang, M., Fuchs, A., Beetle, C., & Kelso, J. A. S. (2020). Coordination Dynamics: A Foundation for Understanding Social Behavior. *Frontiers in Human Neuroscience*, *14*, 317.

Toiviainen, P., Burunat, I., Brattico, E., Vuust, P., & Alluri, V. (2020). The chronnectome of musical beat. *NeuroImage*, *216*, 116191.

Tombaugh, T. N., & McIntyre, N. J. (1992). The mini-mental state examination: a comprehensive review. *Journal of the American Geriatrics Society*, *40*(9), 922–935.

Torre, K., & Balasubramaniam, R. (2009). Two different processes for sensorimotor synchronization in continuous and discontinuous rhythmic movements. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, *199*(2), 157–166.

Triplett, J. W., Owens, M. T., Yamada, J., Lemke, G., Cang, J., Stryker, M. P., & Feldheim, D. A. (2009). Retinal input instructs alignment of visual topographic maps. *Cell*, *139*(1), 175–185.

Tsakiris, M. (2010). My body in the brain: a neurocognitive model of body-ownership. *Neuropsychologia*, *48*(3), 703–712.

Tsakiris, M. (2017). The multisensory basis of the self: From body to identity to others. *The Quarterly Journal of Experimental Psychology*, *70*(4), 597–609.

Tsakiris, M., & Haggard, P. (2005). The rubber hand illusion revisited: visuotactile integration and self-attribution. *Journal of Experimental Psychology. Human Perception and Performance*, *31*(1), 80–91.

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

Tuller, B., & Kelso, J. A. (1989). Environmentally-specified patterns of movement coordination in normal and split-brain subjects. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, *75*(2), 306–316.

Tuller, B., & Kelso, J. A. S. (1985). Coordination in normal and split-brain patients. *Psychonomic Society, Boston, Mass*.

Turvey, M. T. (1992). *Ecological foundations of cognition: Invariants of perception and action*. Retrieved from https://psycnet.apa.org/record/1992-97909-004

Tzanaki, P. (2022). The Positive Feedback Loop of Empathy and Interpersonal Synchronisation: Discussing a Theoretical Model and its Implications for Musical and Social Development. *Music & Science*, *5*, 20592043221142716.

Van den Berghe, P., Lorenzoni, V., Derie, R., Six, J., Gerlo, J., Leman, M., & De Clercq, D. (2021). Music-based biofeedback to reduce tibial shock in over-ground running: a proof-of-concept study. *Scientific Reports*, *11*(1), 4091.

van der Meer, L., Groenewold, N. A., Nolen, W. A., Pijnenborg, M., & Aleman, A. (2011). Inhibit yourself and understand the other: neural basis of distinct processes underlying Theory of Mind. *NeuroImage*, *56*(4), 2364–2374.

van der Steen, J. T., Smaling, H. J., van der Wouden, J. C., Bruinsma, M. S., Scholten, R. J., & Vink, A. C. (2018). Music-based therapeutic interventions for people with dementia. *Cochrane Database of Systematic Reviews* , *7*(7), CD003477.

van der Steen, M. C. M., & Keller, P. E. (2013). The ADaptation and Anticipation Model (ADAM) of sensorimotor synchronization. *Frontiers in Human Neuroscience*, *7*, 253.

Van Dyck, E., Buhmann, J., & Lorenzoni, V. (2021). Instructed versus spontaneous entrainment of running cadence to music tempo. *Annals of the New York Academy of Sciences*, *1489*(1), 91–102.

Van Dyck, E., Moens, B., Buhmann, J., Demey, M., Coorevits, E., Dalla Bella, S., & Leman, M. (2015). Spontaneous Entrainment of Running Cadence to Music Tempo. *Sports Medicine - Open*, *1*(1), 15.

Van Ede, F., De Lange, F., Jensen, O., & Maris, E. (2011). Orienting attention to an upcoming tactile event involves a spatially and temporally specific modulation of sensorimotor alpha-and beta-band oscillations. *Journal of Neuroscience*, *31*(6), 2016–2024.

van Ede, F., Jensen, O., & Maris, E. (2010). Tactile expectation modulates pre-stimulus beta-band oscillations in human sensorimotor cortex. *NeuroImage*, *51*(2), 867–876.

van Ede, F., Quinn, A. J., Woolrich, M. W., & Nobre, A. C. (2018). Neural Oscillations: Sustained Rhythms or Transient Burst-Events? *Trends in Neurosciences*, *41*(7), 415–417.

Van Kerrebroeck, B. (2023). *Extended musical spaces: participatory and creative music making in digitally augmented environments*. Ghent University.

Van Kerrebroeck, B., & Maes, P.-J. (2021). A Breathing Sonification System to Reduce Stress During the COVID-19 Pandemic. *Frontiers in Psychology*, *12*, 623110.

Van Kets, N., Moens, B., Bombeke, K., Durnez, W., Maes, P.-J., Van Wallendael, G., … Lambert, P. (2021). Art and Science Interaction Lab -- A highly flexible and modular interaction science research facility. Retrieved from http://arxiv.org/abs/2101.11691

van Noorden, L., & Moelants, D. (1999). Resonance in the Perception of Musical Pulse. *Journal of New Music Research*, *28*(1), 43–66.

van Pelt, S., Heil, L., Kwisthout, J., Ondobaka, S., van Rooij, I., & Bekkering, H. (2016). Beta- and gamma-band activity reflect predictive coding in the processing of causal events. *Social Cognitive and Affective Neuroscience*, *11*(6), 973–980.

van Ulzen, N. R., Lamoth, C. J. C., Daffertshofer, A., Semin, G. R., & Beek, P. J. (2008). Characteristics of instructed and uninstructed interpersonal coordination while walking side-by-side. *Neuroscience Letters*, *432*(2), 88–93.

Varlet, M., Marin, L., Issartel, J., Schmidt, R. C., & Bardy, B. G. (2012). Continuity of visual and auditory rhythms influences sensorimotor coordination. *PloS One*, *7*(9), e44082.

Varlet, M., Marin, L., Raffard, S., Schmidt, R. C., Capdevielle, D., Boulenger, J.-P., … Bardy, B. G. (2012). Impairments of social motor coordination in schizophrenia. *PloS One*, *7*(1), e29772.

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), 116303.

Ventre-Dominey, J., Gibert, G., Bosse-Platiere, M., Farnè, A., Dominey, P. F., & Pavani, F. (2019). Embodiment into a robot increases its acceptability. *Scientific Reports*, *9*(1), 10083.

Vialatte, F.-B., Maurice, M., Dauwels, J., & Cichocki, A. (2010). Steady-state visually evoked potentials: focus on essential paradigms and future perspectives. *Progress in Neurobiology*, *90*(4), 418–438.

Vidal, M., Rosso, M., & Aguilera, A. M. (2021). Bi-smoothed functional independent component analysis for eeg artifact removal. *Science in China, Series A: Mathematics*. Retrieved from https://www.mdpi.com/2227-7390/9/11/1243

von der Lühe, T., Manera, V., Barisic, I., Becchio, C., Vogeley, K., & Schilbach, L. (2016). Interpersonal predictive coding, not action perception, is impaired in autism. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *371*(1693), 20150373.

Von Holst, E. (1973). *The behavioural physiology of animals: the selected papers of Erich von Holst*. Methuen.

von Schnehen, A., Hobeika, L., Huvent-Grelle, D., & Samson, S. (2022). Sensorimotor Synchronization in Healthy Aging and Neurocognitive Disorders. *Frontiers in Psychology*, *13*, 838511.

Vroomen, J., Bertelson, P., & de Gelder, B. (2001). Directing spatial attention towards the illusory location of a ventriloquized sound. *Acta Psychologica*, *108*(1), 21–33.

Vuust, P., Dietz, M. J., Witek, M., & Kringelbach, M. L. (2018). Now you hear it: a predictive coding model for understanding rhythmic incongruity. *Annals of the New York Academy of Sciences*. doi:10.1111/nyas.13622

Vuust, P., Heggli, O. A., Friston, K. J., & Kringelbach, M. L. (2022). Music in the brain. *Nature Reviews. Neuroscience*, *23*(5), 287–305.

Vuust, P., & Witek, M. A. G. (2014). Rhythmic complexity and predictive coding: a novel approach to modeling rhythm and meter perception in music. *Frontiers in Psychology*, *5*, 1111.

Wachowiak, M. (2011). All in a Sniff: Olfaction as a Model for Active Sensing. *Neuron*, *71*(6), 962–973.

Wagner, M. J., & Smith, M. A. (2008). Shared internal models for feedforward and feedback control. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *28*(42), 10663–10673.

Washburn, A., Kallen, R., Lamb, M., Stepp, N., Shockley, K. D., & Richardson, M. J. (2017). Anticipatory Synchronization in Artificial Agents. *CogSci*. Retrieved from https://www.researchgate.net/profile/Michael-Richardson-21/publication/321171739_Anticipatory_Synchronization_in_Artificial_Agents/lin

ks/5a136e304585158aa3e62b05/Anticipatory-Synchronization-in-Artificial-Agents.pdf

Webber, C. L., Jr, & Zbilut, J. P. (1994). Dynamical assessment of physiological systems and states using recurrence plot strategies. *Journal of Applied Physiology*, *76*(2), 965–973.

Wilson, M., & Cook, P. F. (2016). Rhythmic entrainment: Why humans want to, fireflies can't help it, pet birds try, and sea lions have to be bribed. *Psychonomic Bulletin & Review*, *23*(6), 1647–1659.

Witek, M. A. G., Clarke, E. F., Wallentin, M., Kringelbach, M. L., & Vuust, P. (2015). Correction: Syncopation, Body-Movement and Pleasure in Groove Music. *PloS One*, *10*(9), e0139409.

Witz, K., Hinkle, D. E., Wiersma, W., & Jurs, S. G. (1990). Applied statistics for the behavioral sciences. *Journal of Educational Statistics*, *15*(1), 84.

Wobbrock, J. O., Findlater, L., Gergle, D., & Higgins, J. J. (2011, May 7). The aligned rank transform for nonparametric factorial analyses using only anova procedures. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 143–146. Presented at the Vancouver, BC, Canada. New York, NY, USA: Association for Computing Machinery.

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *358*(1431), 593–602.

Wraga, M., Shephard, J. M., Church, J. A., Inati, S., & Kosslyn, S. M. (2005). Imagined rotations of self versus objects: an fMRI study. *Neuropsychologia*, *43*(9), 1351–1361.

Xenides, D., Vlachos, D. S., & Simos, T. E. (2008). Synchronization in complex systems following a decision based queuing process: rhythmic applause as a test case. *Journal of Statistical Mechanics* , *2008*(07), P07017.

Yamanishi, J., Kawato, M., & Suzuki, R. (1980). Two coupled oscillators as a model for the coordinated finger tapping by both hands. *Biological Cybernetics*, *37*(4), 219–225.

Yoo, G. E., & Kim, S. J. (2016). Rhythmic Auditory Cueing in Motor Rehabilitation for Stroke Patients: Systematic Review and Meta-Analysis. *Journal of Music Therapy*, *53*(2), 149–177.

Zacks, J. M., & Michelon, P. (2005). Transformations of visuospatial images. *Behavioral and Cognitive Neuroscience Reviews*, *4*(2), 96–118.

Zalta, A., Petkoski, S., & Morillon, B. (2020). Natural rhythms of periodic temporal attention. *Nature Communications*, *11*(1), 1051.

Zamm, A., Loehr, J. D., Vesper, C., Konvalinka, I., Kappel, S. L., Heggli, O. A., … Keller, P. E. (2023). *A Practical Guide to EEG Hyperscanning in Joint Action Research: From Motivation to Implementation*. Retrieved from https://psyarxiv.com/fy4kn/download?format=pdf

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: auditory–motor interactions in music perception and production. *Nature Reviews. Neuroscience*, *8*(7), 547–558.

Zbilut, J. P., & Webber, C. L. (1992). Embeddings and delays as derived from quantification of recurrence plots. *Physics Letters. A*, *171*(3), 199–203.

Zeng, H., & Sanes, J. R. (2017). Neuronal cell-type classification: challenges, opportunities and the path forward. *Nature Reviews. Neuroscience*, *18*(9), 530–546.

Zhang, M., Kelso, J. A. S., & Tognoli, E. (2018). Critical diversity: Divided or united states of social coordination. *PloS One*, *13*(4), e0193843.

Zhou, B., Yang, S., Mao, L., & Han, S. (2014). Visual feature processing in the early visual cortex affects duration perception. *Journal of Experimental Psychology. General*, *143*(5), 1893–1902.

Zhou, G., Bourguignon, M., Parkkonen, L., & Hari, R. (2016). Neural signatures of hand kinematics in leaders vs. followers: A dual-MEG study. *NeuroImage*, *125*, 731–738.

Zivotofsky, A. Z., & Hausdorff, J. M. (2007). The sensory feedback mechanisms enabling couples to walk synchronously: an initial investigation. *Journal of Neuroengineering and Rehabilitation*, *4*, 28.

Zoefel, B., & VanRullen, R. (2015). Selective perceptual phase entrainment to speech rhythm in the absence of spectral energy fluctuations. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(5), 1954–1964.