University of Genova

PhD Program in Bioengineering and Robotics

Coordination mechanisms in joint action

by

Cecilia De Vicariis

Thesis submitted for the degree of *Doctor of Philosophy* (35° cycle)

May 2023

Vittorio Sanguineti	Supervisor
Paolo Massobrio	Head of the PhD program

Thesis Jury:

Andrea d'Avella, *University of Messina, Italy*	External examiner
Strahinja Dosen, *Aalborg University, Denmark*	External examiner
Silvio Paolo Sabatini, *University of Genoa, Italy*	Internal examiner

Dibris

Department of Informatics, Bioengineering, Robotics and Systems Engineering

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 65,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures.

Cecilia De Vicariis
May 2023

Acknowledgements

First and foremost I am extremely grateful to my supervisor, Prof. Vittorio Sanguineti, for his continuous support and patience during my PhD study and in the present times. He transmitted invaluable knowledge and has been capable of keeping me on track and encouraged me all the time in my academic research and in particular, during the pandemic.

I am thankful to Prof. Etienne Burdet, Prof. Ekaterina Ivanova, and the Human Robotics Group which hosted me for the last six months of my PhD. I am grateful I had the opportunity to collaborate with them and I hope we will continue to investigate interesting issues together in the future.

I would like to thank Dr. Vinil T. Chacochan, whose previous work has set up a solid starting point for my research. I am deeply thankful also to Dr. Giulia Pusceddu, which proactively collaborated in the artificial partner development as part of her master's thesis.

I am also grateful to my internal reviewers, Prof. Andrea Canessa and Prof. Maura Casadio, for the advice which came along this journey. A particular thanks go to Prof. Maura Casadio, which introduced me to the research and scientific world from my bachelor thesis and supported me with precious advice after my master's studies.

I would also like to thank two of our department's most important people, Giorgio Carlini and Filippo Sante, for their help with the technical aspects of my work which I could have not carried out without their teaching and support.

I thank those professors or researchers who contributed with an opinion, a discussion, or some words of comfort and reassurance.

I am deeply grateful to all my fellow lab-mates. I am very happy I shared this period of my life with such amazing people. You not only contributed to shaping and improving our work environment by continuously supporting each other, but you also become a group of friends outside laboratories and offices.

I am grateful to the incredible friends ('Most Beautiful' and all the others) a person could ever wish for, with the incredible capacity to spread joy and make every day special, their invaluable companionship, and encouragement whenever was needed.

A huge thanks go to my family who always supported and reassured me, encouraging me to fight for what I want without constraints.

And finally, I am grateful to Andrea who supports me and respects my difficult decisions all the time.

Abstract

Coordinated actions of multiple agents that bring about a change in the environment are referred as ‘joint action’. Joint action is ubiquitous in our daily life - for example while carrying objects, playing sports, making music, or during rehabilitation. Many of them - for instance, walking in a narrow passage without bumping into one another - are not planned in advance. In particular, when two humans physically interact, sensory consequences and motor outcomes are not entirely predictable as they also depend on the other’s agent actions. It may be that agents predict partners’ actions and their outcomes as well. However, it is not clear whether and to what extent we actually do this. While sensorimotor mechanisms underlying individual actions have been widely investigated and many foundational principles have been identified, a clear picture of how joint action is controlled and develops over time is still missing.

This work intends to investigate the perceptual and control mechanisms underlying joint action, with a particular focus on emergent coordination in physical interaction. We explore scenarios in which interacting agents need to negotiate a shared strategy through the haptic channel. To this end, we used a combination of computational and experimental approaches. We specifically focused on a few aspects: (i) the subjective experience of control during joint action, (ii) the translation of a computational model of joint action into an architecture that can be used to develop bio-inspired artificial agents, and (iii) learning and decision-making in joint action.

We experimentally addressed physical human-human interaction using a dyadic robotic setup. It consists of a pair of haptic interfaces which can be programmed to implement different types of interaction. We used tasks where players were mechanically coupled and had to perform goal-directed movements in different interaction modalities - sensory modalities and tasks.

When addressing joint action it is important to account for the more subjective experiences and feelings which may emerge and affect performances in joint action. To this end, we extended a previously proposed computational model of interaction to describe the sense of agency during an interaction. Sense of agency refers to the subjective experience of control

we have over our actions and their outcomes. We provided a general framework which reconciles more deterministic accounts of action with the more subjective experiences of control. We proposed that the sense of agency emerges within the sensorimotor control loop and that experiences of self and joint agency coexist during interactions. These are weighted depending on the interactive scenario and actively affect human behavior.

Investigating the mechanisms underlying joint action is crucial per se to understand humans behaviors underpinnings. However, these findings may serve the development of biomimetic artificial agents. We used the interaction model to devise an artificial partner architecture that is capable of estimating human action and establishing coordination. This architecture is a promising tool that may be used to further investigate joint action in a more controllable way but also to develop artificial agents which may be used in a variety of applicative scenarios - e.g. automation, and neuro-rehabilitation.

When people share the same final goal but have incompatible sub-goals deciding to collaborate does not depend on the information available about the partner. To disentangle these findings we developed a computational model of interaction based on game theory and optimal Bayesian estimation. This can be used to make predictions and analyze experimental data, characterizing the interactive participants. We found that other than sensory reliability, initial attitude toward exploration leads dyads to establish coordination. We also investigated how priming interacting agents with identical or different cues affects interaction and the establishment of coordination in complex arm movements. We observed that people tend to synchronize their motions as the mechanical coupling is activated, however, agents tended to use their preferred initial solution which involves some degree of effort instead of negotiating a shared path. This provides insights into understanding to what extent interacting agents represent the partner and in which measure they use such information to select their own action.

Overall, this thesis work attempted to investigate joint action by investigating different aspects of joint action. We not only focused on a mechanistic description of physical human-human interaction, but we also accounted for the more subjective experience of action and suggested an architecture that translates insights from the study of human-human interaction into the development of an application which can be used to control artificial agents which are capable of interacting with humans in a natural and intuitive way. We think that this approach is crucial when addressing a multifaceted and complicated phenomenon as joint action.

Table of contents

List of figures **xi**

List of tables **xv**

1 Introduction **1**

1.1 Negotiation in physical human-human interaction 3
1.2 The subjective experience of control in joint action 5
1.3 Understanding joint action to develop bio-inspired artificial agents 7
1.4 Aims: motivations and objectives 9
1.5 Outline 9

2 Game theory and partner representation in joint action: toward a computational theory of joint agency **13**

2.1 Introduction 14
2.2 Computational Model of joint agency 16
  2.2.1 General framework 16
  2.2.2 From self agency to joint agency 19
2.3 Joint agency in a sensorimotor interactive task 28
  2.3.1 Task 29
  2.3.2 Model Implementation 29
  2.3.3 Simulation results 30
2.4 Discussion 33
  2.4.1 Model prediction 34
  2.4.2 Learning in joint action 37
  2.4.3 Relation with previous models of joint action 37
2.5 Conclusion 40

3 Artificial Partners to Understand Joint Action: Representing Others to Develop Effective Coordination 49
3.1 Introduction 50
3.2 Modeling framework 51
  3.2.1 Optimality in perception and action 51
  3.2.2 Partner representation 52
  3.2.3 Game theory and fictitious play 52
3.3 Materials and Methods 54
  3.3.1 Artificial Partner Architecture 54
  3.3.2 Experiments 60
  3.3.3 Data Analysis 64
3.4 Results 65
  3.4.1 Experiment 1 - Partner model 65
  3.4.2 Experiment 2 - Learning coordination 66
3.5 Discussion 69
  3.5.1 The artificial partner estimates the actions of its opponent 70
  3.5.2 Artificial and human partner learn to coordinate 70
  3.5.3 A general platform to study joint action 71
  3.5.4 Artificial partners as diagnostic and rehabilitation tools 71
3.6 Conclusions 72

4 Computational joint action: dynamical models to understand the development of joint coordination 73
4.1 Introduction 73
4.2 General Computational Framework 75
  4.2.1 Joint tasks as quadratic games 76
  4.2.2 Sensory system 77
  4.2.3 Partner model 78
  4.2.4 Action selection and strategy choice 79
  4.2.5 Model identification 81
4.3 Applications to joint scenarios 83
  4.3.1 Prisoner's dilemma 83
  4.3.2 Stag Hunt 85
  4.3.3 2-VP task 87
4.4 Results 89

4.4.1 Prisoner's Dilemma 89
4.4.2 Stag Hunt 91
4.4.3 2-VP task 92
4.5 Discussion 93
4.6 Conclusion 95

5 Learning and partner representation in joint coordination 97
5.1 Introduction 97
5.2 Materials and methods 99
5.2.1 Experimental Apparatus and Task 99
5.2.2 Subjects 101
5.2.3 Data analysis 102
5.3 Results 106
5.3.1 Experiments 106
5.4 Discussion 113

6 Haptic communication in redundant sensorimotor tasks 115
6.1 Introduction 115
6.2 Materials and Methods 117
6.2.1 Experimental apparatus and task 117
6.2.2 Experiment 1 118
6.2.3 Experiment 2 120
6.2.4 Data Analysis 122
6.3 Results 124
6.3.1 Experiment 1 - Collaboration 124
6.3.2 Experiment 2 - Negotiation 129
6.4 Discussion 134

7 Conclusions and outlook 141
7.1 Major contributions 141
7.1.1 The sense of agency in joint action 141
7.1.2 From human-human interaction to human-artificial agent interaction 142
7.1.3 Negotiation through haptic communication 142
7.2 General conclusion 143
7.3 List of Publications and Conferences 146

References **149**

Appendix A Supplementary figures - Chapter 4 **161**

Appendix B Supplementary figures - Chapter 5 **163**

Appendix C Supplementary figures - Chapter 6 **167**

C.1 Experiment 1 - Collaboration 167

C.2 Experiment 2 - Negotiation 172

List of figures

1.1 Joint action building blocks 3
1.2 Comparator model 6
1.3 Aritficial Partner and Human Dynamic Clamp paradigms 8

2.1 Computational model of self agency 18
2.2 Computational model of joint agency 22
2.3 Joint state observer 23
2.4 A joint action scenario and its game theoretic interpretation 24
2.5 Between-trials simulation results 32
2.6 Within trial simulation results 33

3.1 Artificial partner architecture 53
3.2 Two-via point joint action task 60
3.3 Experiment 1 - Partner model 63
3.4 Experiment 2 - Coordination 66
3.5 Trajectories and speed profiles 68

4.1 Computational model 76
4.2 Computational model, extended to account for multiple strategy choices . . 77
4.3 Prisoner dilemma motor game 84
4.4 A continuous version of the Stag Hunt game 86
4.5 The 2-VP task 88
4.6 Prisoner dilemma simulation results 90
4.7 Stag Hunt simulation results 91
4.8 Simulation results 93
4.9 Experimental Results 94
4.10 Fitting results 94

5.1 Experimental apparatus and protocol 99
5.2 Nash equilibria and computational model 105
5.3 Coordination behavior 107
5.4 Initial and Final Synchronization. 109
5.5 Synchronization at via-points 109
5.6 Action selection 110
5.7 Estimated model parameters 112
5.8 Determinants of learning 112

6.1 Experiment 1 119
6.2 Experiment 2 121
6.3 Trajectories and speed profiles 125
6.4 Asynchrony 126
6.5 Dynamic Time Warping 127
6.6 Speed correlation 128
6.7 Interaction Force 128
6.8 Trajectories and speed profiles 130
6.9 Asynchrony 131
6.10 Dynamic Time Warping 132
6.11 Speed correlation 133
6.12 Interaction force 134
6.13 Software architecture and dyadic setup 137

7.1 From human-human interaction to artificial therapists for neurorehabilitation 145

A.1 Model parameters 162

B.1 Trajectories for H group 164
B.2 Trajectories for VH group 165

C.1 Trajectories and speed profiles of dyad 1 168
C.2 Trajectories and speed profiles of dyad 2 169
C.3 Trajectories and speed profiles of dyad 3 170
C.4 Trajectories and speed profiles of dyad 4 171
C.5 Trajectories and speed profiles of dyad 5 172
C.6 Trajectories and speed profiles of dyad 1 173
C.7 Trajectories and speed profiles of dyad 2 173

C.8 Trajectories and speed profiles of dyad 3 174
C.9 Trajectories and speed profiles of dyad 4 174
C.10 Trajectories and speed profiles of dyad 5 175
C.11 Trajectories and speed profiles of dyad 6 175
C.12 Trajectories and speed profiles of dyad 7 176

List of tables

4.1 Summary of model parameters 81

6.1 Statistical analysis for Experiment 1 124
6.2 Statistical analysis for Experiment 2 129

Chapter 1

Introduction

Humans are natively social beings as from infancy our behaviors are influenced by verbal, visual, and haptic cues from other people. Many of our everyday activities occur in social settings and are coordinated with others. Even seemingly simple and natural interactions, like a pair of workers sawing timber with back and forth movements, a mother helping her child to walk, a couple on a dance floor, two children playing a ball game, passing someone walking toward you in a narrow alley or providing physical therapy to a patient, require that two individual minds are somehow connected and their bodies coordinated (Sebanz et al., 2006). Performing an action with someone requires predicting what others are going to do next, adjusting our behavior, and achieving precise temporal coordination by taking those predictions into account (Vesper et al., 2010).

How do interactions with others change our behavior? Many of those interactions require active coordination, which is manifested by physical and cognitive responses related to explicit knowledge of the interacting partner. Some studies stressed the role of explicit reasoning about the partner and the interaction. Language (Bacharach, 1999; Clark, 1996), attribution of intention (Carruthers and Smith, 1996) and commitments (Clark, 2020) have crucial roles in the achievement of coordination in this framework. However, coordination necessarily relies on direct perception-action link (Schmidt and Richardson, 2008; Tognoli et al., 2020) and more subtle and implicit mechanisms. Knoblich et al. (2011) proposed a distinction between planned and emergent coordination. Planned coordination refers to those situations in which individuals explicitly agree on a shared plan to achieve a joint outcome. These interactions are driven by the development of the representations of the joint goal and own individual contributions to the desired outcome. Emergent coordination occurs between individuals who have no plans to act jointly, but also during joint actions. As such, emergent coordination is pervasive in our daily lives.

Efforts to depict a minimal unified architecture of the mechanisms underlying coordination has been done by Vesper et al. (2010). Vesper et al. (2016, 2010) defined the basic building blocks for efficient joint coordination - i.e. representations, processes, and coordination smoothers - see Figure 1.1. Internal representations need to account for the joint setting and the interacting agents. Individuals in a joint action have to build internal representations of their tasks and goal. It is not necessary, although typically very useful, to represent the other's task as well. When acting together with others, representing a co-actor's task can be beneficial as it may facilitate agents in predicting others' actions and integrating them into their action plans. It can be argued that agents do not need to predict and know others' tasks to predict their partner's actions. The nature and the level of detail of the representations about a partner is an open issue. Further, Vesper et al. (2010) proposed that two processes manage the outlined representations. Monitoring refers to the continuous observation of whether actions are unfolding as specified by task or goal representations. Joint action may require monitoring the joint outcomes of interacting agents' actions or individual outcomes. Monitoring processes are inherently intermingled with predictive processes. These are paramount as they consist in estimating how actions will unfold based on internal representation and the motor plan of the next action. Predicting the next action outcome based on current goal and task representation and monitoring the course of action to update representations are pivotal for precise coordination. The third category of building blocks proposed by Vesper et al. (2010) is coordination smoothers. Coordination smoothers refer to action modulation strategies that facilitate the achievement of shared coordination. For instance, making our behavior more predictable, reducing movement variability, or making motions more salient, are just two possibilities. Another example is limiting personal action workspace, which conveys information to the partner about where to act to achieve coordination or fulfill a common goal. Also Pezzulo et al. (2013) proposed a signaling theory, underlying the importance of sensorimotor forms of communication to exchange coordination cues.

However, in those frameworks (Pezzulo et al., 2013; Vesper et al., 2010) no attention has been addressed to the decision-making and action-selection processes. When interacting with a partner the number of possible solutions to bring about a certain task increases with respect to individual performances. For this purpose, it is crucial to devise a framework that integrates also the decision-making processes in joint action. These processes have been explained, for individual actions, with the comparator model (Blakemore et al., 2002; Frith et al., 2000) where sensory information and predictions are continuously compared to evaluate the sensory-prediction error. The sensory prediction error is used to update internal

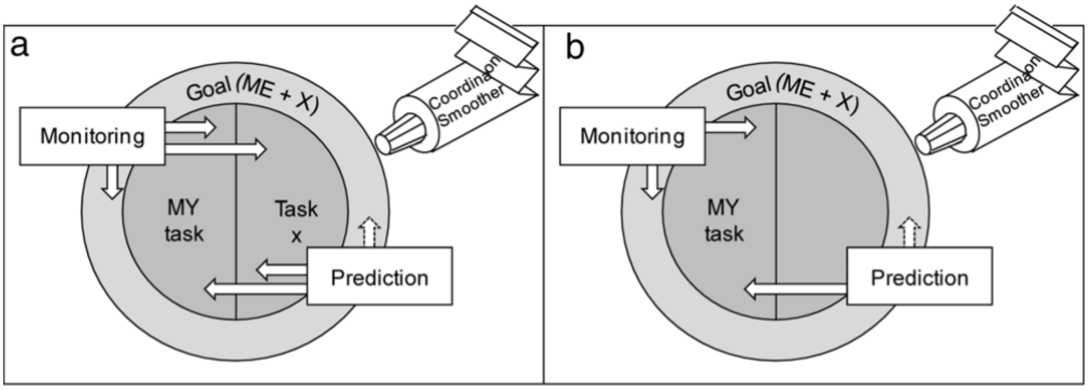

Figure 1.1 Joint action building blocks from Vesper et al. (2010). In the circle, the representations are indicated. The external ring indicates the representation of the joint action goal. On the left, both representations of the own task and the other's task are reported. On the right only own task representation is depicted, to indicate that in joint action is not strictly necessary to represent the partner's task. Processes are reported in the white boxes with arrows toward the representations that are monitored and predicted. The glue stick indicates coordination smoothers.

representations of the environment and the objects involved in the actions and consequently plan and continuously correct our behavior.

1.1 Negotiation in physical human-human interaction

To overcome the complexity of studying joint action in real-life scenarios, researchers often address human-human interaction mediated by machines (Küçüktabak et al., 2021). Peculiar types of joint action are those in which the two partners are physically coupled (Jarrassé et al., 2012; Melendez-Calderon et al., 2015; Reed et al., 2006; Sabu et al., 2020; van der Wel et al., 2011). Object-mediated interaction and more in general physical interaction is enabled by implicit information exchange between interacting individuals using subtle haptic signals within their interaction force (Colomer et al., 2022). These types of interaction can be experimentally addressed using pair of haptic interfaces which can be programmed to implement a variety of interaction modalities (Chackochan and Sanguineti, 2019; Ganesh et al., 2014; Küçüktabak et al., 2021; Melendez-Calderon et al., 2011, 2015; Takagi et al., 2019). This scenario is particularly useful for quantifying various aspects of joint action performance because it allows to experimentally manipulate all aspects of the interaction. Manipulation of robot controllers, interaction modality, and task allows investigating a wide range of joint actions, from collaboration (Avila Mireles et al., 2017; Beckers et al., 2020;

Ganesh et al., 2014; Melendez-Calderon et al., 2015; van der Wel et al., 2011) to competition or scenarios where players may negotiate a shared strategy to achieve their goals (Chackochan and Sanguineti, 2019).

Understanding how humans use the haptic channel as a source of information about the external world and other people is crucial, as we constantly rely on this source of information to interact in real-life settings (Ernst and Banks, 2002). It is widely recognized that humans interact with the environment by minimizing a cost function (Todorov, 2004; Todorov and Jordan, 2002). Cost function minimization models the trade-off between error and effort. In Jarrassé et al. (2012) interacting agents aim at minimizing their effort while performing separate or common actions. Agents receive reward signals and multimodal sensory information. Each agent perceives their error and estimates their partner's error. Overall, the interaction behavior results from the combination of the minimization of the individual cost functions. However, it can be argued that to efficiently interact with a partner we do not need to represent the partner's error and the partner's effort. More in general, it is plausible that in the cost function of at least one of the interactive agents appears a term that account for the interaction.

Game theory has been proposed as a computational framework to extend optimal control (Braun et al., 2009, 2011; Chackochan and Sanguineti, 2019; Grau-Moya et al., 2013; Jarrassé et al., 2012; Yoshida et al., 2008). Game theory provides the analytic and computational substrate for the decision-making and control processes underlying interactive tasks, where agents' activities affect each other. A 'game' is a situation 'involving two or more individuals whose interests are neither completely opposed nor completely coincident' (Nash, 1953). Cooperative games are those scenarios where interactive agents explicitly agree on a common strategy and behave as a collective (Bacharach, 1999; Newton, 2018; Tuomela, 2007). In other scenarios, there is no prior explicit agreement and we may not be equally aware of the goals of our partner as we are of our own. Situations in which each player autonomously decides the action to take are referred to as 'non-cooperative games'.

Nash equilibrium refers to a situation where none of the agents can unilaterally improve their benefit (Nash, 1951, 1953). This concept is one of the most central solution concepts to studying strategic interactions between multiple players. It has been shown to capture sensorimotor interactions between players that are haptically coupled (Braun et al., 2009, 2011; Chackochan and Sanguineti, 2019; Jarrassé et al., 2012; Li et al., 2019). More recently, Lindig-León et al. (2021a) suggested that bounded rational response equilibria provide a general tool to explain sensorimotor interactions that include the Nash equilibrium as a special case with perfect information about the partner (Harsanyi, 1967). In this framework, players are bounded rational, lacking perfect rationality to maximize their expected utility (Goeree and Holt, 1999; Rubinstein, 1998; Von Neumann and Morgenstern, 1944). This may account for scenarios where agents may not know all possible outcomes or the utility functions of the other players, they may have incomplete knowledge, model uncertainty, or lack computational resources. Quantal response equilibria (QRE) have been used to address deviation from Nash equilibria (McKelvey and Palfrey, 1995, 1998). The idea underlying the definition of QRE is that players make infinitesimal errors during their action selection. In this perspective, best response functions become probabilistic rather than deterministic. Using a probabilistic framework allows us to inherently account for human variability and it is well suitable to address the more subjective experiences of interaction.

Overall, game theory allows addressing in systematic and rigorous mathematical formalisms several aspects of human-human interaction. Social interaction types can be modeled depending on the cost and benefits of each interacting agent and on the information available. Lastly, games with repeated play can be used to address learning, and games' solutions can describe equilibria situations in joint action.

1.2 The subjective experience of control in joint action

Understanding how human behavior is influenced by others is intrinsically and inevitably related to the more subjective experiences of action. Our perception of the task, the value we attribute to the task, and our personal attitudes toward action are different when performing alone or with other people (Pacherie, 2014). Consequently action selection processes - e.g. decision-making, and planning - are inherently different. The sense of agency is the subjective experience of being in control of our own actions and through them of events in the external world. This feeling has been mainly addressed in individual actions. It is granted that the sense of agency relies on cues related to the sensorimotor level of action and cues related to higher-order cognitive processes (Moore and Fletcher, 2012; Pacherie, 2008; Synofzik et al., 2008; Zalla and Sperduti, 2015). Further, the principle of congruence underlies the emergence of the sense of agency. The comparator model - see Figure 1.2 - has been proposed to explain sensorimotor control and the sense of agency which results from the match between intentions, predictions, and observations is pivotal to the experience of control over an action (Blakemore et al., 2002; Frith et al., 2000).

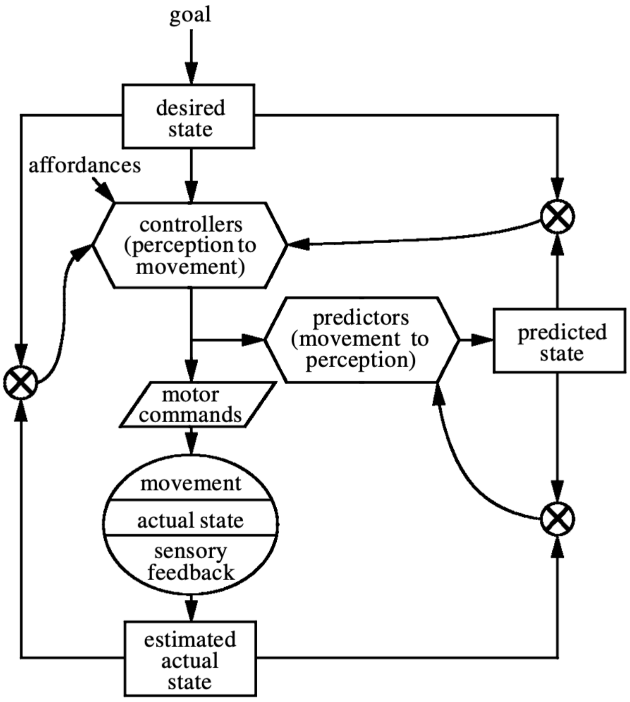

Figure 1.2 Comparator model from Frith et al. (2000).

Less attention has been devoted to the sense of agency in interactive settings. During joint action, it is likely that a sense of joint agency emerges. This experience, extending previous definitions, can be regarded as the sense of shared control over a joint action and its consequences on the environment, and likely depends on the matching of joint predictions with the actual action outcomes. However, little is known about how and to what extent we understand and account for partners during an interaction and how this affects our subjective experiences of action. It has been suggested that the sense of self and joint agency coexist during joint action (Bolt and Loehr, 2017; Bolt et al., 2016) and are modulated by different cues (Le Bars et al., 2020). Theoretical and philosophical accounts of agency suggested that we may experience two types of sense of joint agency. Depending on the interactive setting, we may experience a sense of 'we-agency' where the joint agency emerges at the expense of the sense of self-agency (Gallotti and Frith, 2013; van der Wel, 2015) or a sense of 'shared agency' where self and joint agency grow in parallel (Bolt and Loehr, 2017; Dewey et al., 2014; Le Bars et al., 2020). However, a clear general picture of the mechanisms underlying the subjective experience of control in joint action is still missing.

Understanding the mechanisms underlying the sense of agency in interactive settings may help to disentangle pathologies and disturbs where social interaction is impaired (Blakemore et al., 2002; Frith et al., 2000). Autism spectrum disorder (Zalla and Sperduti, 2015) and schizophrenia (Izawa et al., 2016) are just two examples of pathological conditions which may benefit from the understanding of the interaction processes. Further, with the surge of artificial and robotics agents, it is important to assess the subjective experiences of actions in human-human settings to depict a reference for the assessment of the experiences in interactions with artificial agents.

1.3 Understanding joint action to develop bio-inspired artificial agents

Investigating how humans interact among them is crucial to fill the gaps that are still open in the study of motor control in healthy and pathological contexts. However, when studying joint action, it is difficult to experimentally disentangle the individual contributions and the way each participant reacts to their partner's actions.

A way to overcome this problem is to design experiments where human participants interact with simulated, human-like artificial partners (APs), whose personal traits and interaction strategies can be manipulated experimentally. This paradigm has been inspired by the dynamic patch clamp used in neurophysiology (Prinz et al., 2004; Sharp et al., 1993a,b), where artificial membrane and synaptic conductances are simulated and implemented into biological neurons to understand how these conductances combine and interact to produce neuronal and neural behaviors. This approach is often referred to as Virtual Partner Interaction (VPI) (Kelso et al., 2009) or Human Dynamic Clamp (HDC) (Dumas et al., 2014a) - see Figure 1.3. The original HDC approach relied on a theoretical model of coordination dynamics which involves coupled non-linear oscillators (Haken et al., 1985). In particular, they developed a model of a virtual partner with human-like behavior and made it interact with real- people. In other words, they studied real-time bidirectional interaction between humans and a robot controlled through an empirically grounded computational model of interaction. They set up an experiment where the human subject was asked to flex and extend their index finger initially following the rhythm of an auditory metronome at 1 Hz while looking at a blank screen. After some time, the virtual partner appeared on the screen, and the human was asked to accelerate their pace and to stop at will. The VP, after an initial lag, entered into in-phase coordination with the human. The HDC paradigm provided important contributions to the study of discrete joint action. However, crucial determinants

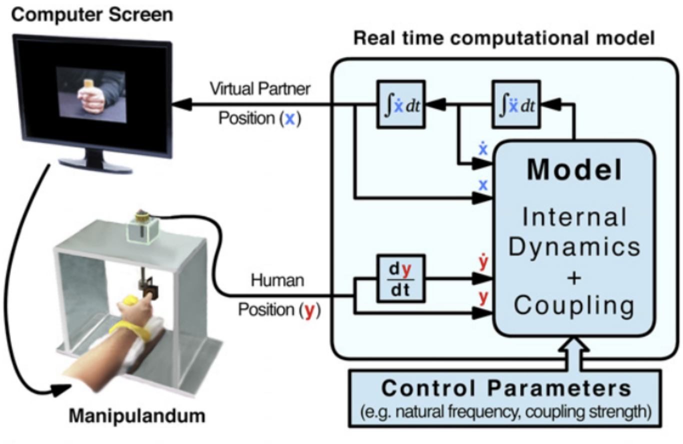

Figure 1.3 Aritficial Partner (Kelso et al., 2009) and Human Dynamic Clamp (Dumas et al., 2014a) paradigms.

of sensorimotor interaction, like perception, decision-making, and control mechanisms need to be explicitly modeled to develop more versatile artificial agents.

Another major challenge for joint action research lies in its possible contribution to the development of interactive artificial agents. We are every day more involved in interaction with novel technological tools which help us in our everyday activities. Translating the understanding of the mechanisms underlying human-human interaction into the development of artificial agents and tools which can be intuitively and naturally integrated into our everyday life is crucial. The challenge is the development of artificial agents that are to be treated as social companions. As such they should evoke the responses and activate the same processes that are typically activated when humans interact with other humans. Inspired by the benefits of human–human interaction, robotic systems have been increasingly used to work in physical interaction with humans. In these contact robots, the physical interaction arises directly, or through an object such as when carrying a large object together. Computational architectures inspired by the study of joint action can be used in automation and manufactory scenarios (Li et al., 2019, 2022) such as microfabrication, or shared driving. The same approach is being proposed in medical contexts, such as surgery and neurorehabilitation (Semprini et al., 2018) to devise personalized treatments and therapies.

1.4 Aims: motivations and objectives

The research work during my PhD was dedicated to providing further understanding of the sensorimotor mechanisms underlying human-human interaction. To this purpose, I used a combination of experimental and computational approaches. The main objective is declined in three specific goals.

- To understand what and to which extent a human represents a partner or an opponent during an interaction, and how the form of this representation affects action performance. To this purpose, we experimentally tested different experimental conditions in sensorimotor interactive tasks by using a novel dual robot experimental setup. We identified indicators of coordination and modeled the underlying decision-making processes to understand which features modulate these mechanisms.
- To investigate the subjective components of action – the sense of agency, i.e. the experience of control - in joint action. This issue is highly related to the first sub-goal. To this end, we proposed a computational model of joint action which predicts how the sense of agency relates to partner representations and action performance.
- To develop an adaptive bio-inspired artificial partner with built-in collaborative capabilities. To this purpose, We developed a general artificial partner architecture and tested experimentally the artificial partner's capabilities to understand the human counterpart and to establish coordination.

Even though the three sub-goals seem to be loosely connected topics, I strongly believe that when addressing joint action none of the three aspects should be neglected. Indeed, it is paramount to observe and empirically explore the mechanisms underlying human-human interaction through experiments and formalize the hypothesis and findings with computational models of interaction. However, we propose that mechanistic views of interaction need to account also for the more subjective aspects of action - i.e. the sense of agency. Further, it is important to recognize the technological relevance of the study of joint action, implying its translation to the development of technological tools which are capable of establishing a natural interaction with humans.

1.5 Outline

In the chapters that follow I describe in detail different studies and works that have been done during the three-years PhD work. Chapter order reflects the chronological order of the studies and activities carried out during my PhD. Due to the pandemic situation and the impossibility to run experiments in the first year, my initial work focused on modeling, then I moved to technological development and these results informed the experimental activities which were mostly concentrated in my last year. Chapter 2 presents a first study addressing the sense of agency in joint action - i.e. the subjective experience of being in control of our actions. It starts off with the introduction of background concepts about the commonly recognized concepts related to sensorimotor control and how these mechanisms need to be extended in joint action. Further, we provide an overview of the sense of agency in individual and joint settings. The chapter continues with the proposal of a computational model of joint action which accounts for the sense of agency. We extended a computational model of interaction Chackochan and Sanguineti (2019) to specifically address the sense of agency in sensorimotor interactive tasks. We discuss model predictions and relate the model to other models of interaction and agency.

In Chapter 3, we present an architecture for the development of artificial agents. We translated an empirically informed interaction model Chackochan and Sanguineti (2019) into a robotic controller. The artificial agent's architecture is embedded, in this case, in a planar haptic interface, and is capable of understanding partner action and using such information to establish a collaboration. We tested the architecture in two experiments and demonstrated that the artificial agent is capable of triggering in the human the same behavior observed in human-human interaction.

Chapter 4 and Chapter 5 address decision-making and learning in joint action. In particular, in Chapter 4 we present a novel computational model of interaction. The computational framework relies on optimal state estimation to predict partner action and uses game theoretic concepts to address learning and decision-making. We demonstrate that the model can be used as a simulation tool to make predictions and as an analytic tool to analyze experiments addressing interaction and to characterize interacting agents in terms of model parameters. In Chapter 5, we report results from an experiment addressing more specifically decision-making when there are multiple strategies to achieve coordination while mechanically coupled. We tested pairs of participants with a dyadic setup involving a couple of 3 DoF haptic interfaces. We analyze experimental data with a set of coordination metrics but we also extend the computational model proposed in Chapter 4 to address this particular setting where people need to decide which strategy to pick and which action to perform.

In Chapter 6 we present preliminary results from two sets of experiments designed and carried out in collaboration with the Human Robotics Group at Imperial College London. In those experiments, we specifically focused on scenarios in which interacting agents are physically coupled and share the same goal but have no constraints in terms of space and time. Further, we tested a setting in which we cued differently the two agents at the beginning of the experiment. This, allowed us to explore the effect of prior individual preferences on the achievement of coordination.

Chapter 7 summarizes and concludes the main results achieved in this thesis and discusses the implications for the field of robot-mediated training. Furthermore, my contributions to the field are explicitly stated and suggestions for future works are given.

Chapter 2

Game theory and partner representation in joint action: toward a computational theory of joint agency

The sense of agency – the subjective feeling of being in control of our own actions – is one central aspect of the phenomenology of action. Computational models provided important contributions toward unveiling the mechanisms underlying the sense of agency in individual action. In particular, the sense of agency is believed to be related to the match between the actual and predicted consequences of our own actions (comparator model). In the study of joint action, models are even more necessary to understand the mechanisms underlying the development of coordination strategies and how the subjective experiences of control emerge during the interaction. In a joint action, we not only need to predict the consequences of our own actions; we also need to predict the actions and intentions of our partner, and to integrate these predictions to infer their joint consequences. Understanding our partner and developing mutually satisfactory coordination strategies are key components of joint action and in the development of the sense of joint agency. Here we discuss a computational architecture which addresses the sense of agency during intentional, real-time joint action. We first reformulate previous accounts of the sense of agency in probabilistic terms, as the combination of prior beliefs about the action goals and constraints, and the likelihood of the predicted movement outcomes. To look at the sense of joint agency, we extend classical

---

Reprinted from: De Vicariis, C., Chackochan, V. T., Sanguineti, V. (2022). Game theory and partner representation in joint action: toward a computational theory of joint agency. Phenomenology and the Cognitive Sciences, 1-30. DOI: 10.1007/s11097-022-09819-5

computational motor control concepts - optimal estimation and optimal control. Regarding estimation, we argue that in joint action the players not only need to predict the consequences of their own actions, but also need to predict partner's actions and intentions (a 'partner model') and to integrate these predictions to infer their joint consequences. As regards action selection, we use differential game theory – in which actions develop in continuous space and time - to formulate the problem of establishing a stable form of coordination and as a natural extension of optimal control to joint action. The resulting model posits two concurrent observer-controller loops, accounting for 'joint' and 'self' action control. The two observers quantify the likelihoods of being in control alone or jointly. Combined with prior beliefs, they provide weighing signals which are used to modulate the 'joint' and 'self' motor commands. We argue that these signals can be interpreted as the subjective sense of joint and self agency. We demonstrate the model predictions by simulating a sensorimotor interactive task where two players are mechanically coupled and are instructed to perform planar movements to reach a shared final target by crossing two differently located intermediate targets. In particular, we explore the relation between self and joint agency and the information available to each player about their partner. The proposed model provides a coherent picture of the inter-relation of prediction, control, and the sense of agency in a broader range of joint actions.

2.1 Introduction

Joint action is ubiquitous in activities of daily living. We coordinate our actions with peers most of our time – for example, while carrying objects, dancing, playing sports, making music or during rehabilitation exercises with a therapist. Researchers have started to investigate how humans deal with joint actions, challenging the assumption traditionally held in cognitive psychology that perception, action and higher-level cognitive processes can be understood by investigating individual minds in isolation. The sense of agency – the subjective feeling of being in control of our own actions and, through them, of events in the external world – has been widely studied in individual action. The sense of agency has been often considered as the outcome of an inference process (Wegner, 2003; Wegner and Wheatley, 1999) which takes place after action completion. Others consider the sense of agency as an immediate component of the (ongoing) perceptual and sensory registration (Blakemore et al., 2002; Frith et al., 2000; Haggard, 2005). Processes which precede the action itself (Chambon et al., 2014) may constitute a 'prospective' component of the sense of agency. During the last decade, a comprehensive theoretical and modeling framework has been developed, which attempts to capture the complexity and the multifaceted nature of the sense of agency (Moore and Fletcher, 2012; Pacherie, 2008; Synofzik et al., 2008; Zalla and Sperduti, 2015). These studies suggest that the sense of self agency results from cues related to the sensorimotor level (efference copy, sensory feedback, sensory predictions) and cues related to higher-order cognitive processes (intentions, beliefs), and relies on one core idea: the principle of congruence – i.e. the match between prior knowledge and intentions, predictions and actual observations. Bayesian concepts have been implicated in the formation of the sense of agency (Fletcher and Frith, 2009; Izawa et al., 2016; Moore and Fletcher, 2012; Synofzik et al., 2013).Comparatively less attention has been addressed to the sense of agency in joint action. A joint action is any activity in which two or more persons need to coordinate their actions or movements in space and time to generate a change in the external world (Sebanz et al., 2006). The sense of joint agency can be regarded as the sense of shared control over a joint action and its consequences on the environment. Pacherie (Pacherie, 2012, 2014) argued that in joint action agents need to predict not only their own actions' outcomes, but also the consequences of the partner's actions. The development of the sense of joint agency likely depends on matching such joint predictions with the actual action outcomes. While information about ourselves is easily accessible, information about the partner typically has a greater degree of uncertainty, leading to possibly incomplete or ambiguous representations. However, little is known about how we perceive other people and to what extent we account for information about our partner during action production and how this contribute to the sense of joint agency. Coordination puts additional requirements in action planning and execution with respect to individual action, thus likely affecting the sense of joint agency. Emergent and planned coordination are regarded as facilitators for the development of the sense of joint agency (Knoblich et al., 2011). Roles distribution and the number of persons involved in the action affect the capability of predicting partner's actions in a joint action (Pacherie, 2014). It is not always straightforward to discriminate contributions in an ongoing action, thus affecting both self and joint agency (Dewey et al., 2014; van der Wel, 2015). Self and joint agency seem to be modulated by different cues (Le Bars et al., 2020), but they coexist in certain situations (Bolt and Loehr, 2017; Bolt et al., 2016). In some forms of joint action, the players may experience a strong sense of joint agency and a weak sense of self agency. This occurs when players perform similar actions with similar effects and synchronous timing, like marching soldiers (Bolt et al., 2016). In other situations, a strong sense of joint agency is accompanied by an equally strong sense of self agency. This is the case when the players are required to produce coordinated yet distinct and complementary actions. To reconcile these findings, it has been suggested that the sense of joint agency comes in two forms, namely we-agency – in which joint agency grows at the expenses of self agency (Gallotti and Frith, 2013; van der Wel, 2015) – and shared-agency – in which both the senses of joint and self agency remain high (Bolt and Loehr, 2017; Dewey et al., 2014; Le Bars et al., 2020). At the very least, these observations suggest that we are capable to assess the senses of joint and self agency separately. Here we propose a general modeling framework to describe how the sense of joint agency emerges from our experience of the goals, ongoing performance and final outcome of our own actions. We take a probabilistic (Bayesian) perspective, which naturally unifies the prospective and retrospective components of agency. We then extend this model to account for joint actions, by also accounting for observations that we maintain distinct senses of self and joint agency, whose relationship is determined by the structural aspects of the interaction – e.g. the interaction modality and the type of joint task. To demonstrate the consistency of our formulation and to highlight its predictions, we use computer simulations to predict the emergence and evolution of the sense of joint agency in a recently reported sensorimotor interaction scenario (Chackochan and Sanguineti, 2019).

2.2 Computational Model of joint agency

2.2.1 General framework

Action generation is the end result of two inter-related processes, i.e. movement control and estimation of the state of the body and the external environment. Movement control is the process of generating motor commands - muscle activations - on the basis of our movement goals and our belief on the current state (position, velocity) of our own body and the external environment (Wolpert, 1997; Wolpert and Miall, 1996). State estimation – or sensorimotor integration – is the process of combining descending motor commands (efferent copy) and sensory information (sensory reafference) to estimate the evolution of the state of the body (and the external environment). Both movement control and state estimation can be understood in terms of optimality principles. There is indeed ample evidence that our nervous system exhibits a close-to-optimal performance in both movement control (Todorov and Jordan, 2002) and sensorimotor integration (Wolpert et al., 1995). Hence the combination of optimal control and optimal estimation is the ideal framework for addressing the sense of agency.

Optimal estimation and agency Action generation is the end result of two inter-related processes, i.e. movement control and estimation of the state of the body and the external environment. Movement control is the process of generating motor commands - muscle activations - on the basis of our movement goals and our belief on the current state (position, velocity) of our own body and the external environment (Wolpert, 1997; Wolpert and Miall, 1996). State estimation – or sensorimotor integration – is the process of combining descending motor commands (efferent copy) and sensory information (sensory reafference) to estimate the evolution of the state of the body (and the external environment). Both movement control and state estimation can be understood in terms of optimality principles. There is indeed ample evidence that our nervous system exhibits a close-to-optimal performance in both movement control (Todorov and Jordan, 2002) and sensorimotor integration (Wolpert et al., 1995). Hence the combination of optimal control and optimal estimation is the ideal framework for addressing the sense of agency.

Optimal control and agency Optimal control posits that humans behave as self-conscious agents, aiming at maximizing their subjective utility, i.e. a trade-off between task-dependent movement cost – in reaching movements, it may simply be the endpoint error – and the perceived effort associated to movement. Hence actions result from a subjective evaluation of their associated costs and/or benefits. This subjective aspect can be summarized by a cost function that incorporates the task goals and requirements. Given a certain goal, a variety of rules – control policies – can be conceived to map body and environment states into motor commands (Wolpert and Kawato, 1998). There is some evidence that the brain selects the optimal motor command as a trade-off between cost and effort (Shadmehr and Krakauer, 2008; Todorov, 2004) – see Figure 2.1. While the role of optimal estimation has been widely acknowledged in the formation of the sense of agency, optimality of control is at least as important. Awareness of what we are doing and how we are trying to achieve our goals is involved in building the sense of agency (Chambon et al., 2014; Moore and Haggard, 2008; Pacherie, 2008). In particular, these multiple cues likely affect our prospective judgement of agency (Lafleur et al., 2020; Synofzik et al., 2013; Zalla and Sperduti, 2015). We may feel to have more or less control over an intended action depending on prior knowledge about the task, the context and the goals to be achieved.

A probabilistic perspective The above considerations point at a comprehensive formulation for the sense of self agency, rooted on probabilistic concepts. Probabilities are intended in subjective or Bayesian terms, i.e. as measures of the degree of belief that some proposition

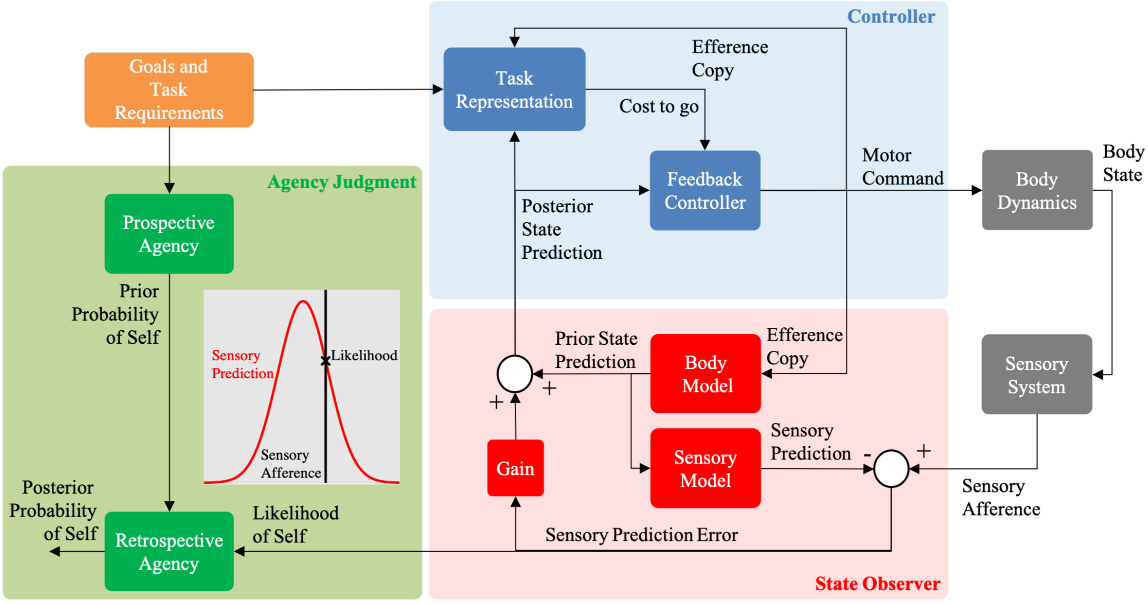

Figure 2.1 Computational model of self agency.The model involves an optimal estimator or state observer block (red) which optimally combines sensory information and motor command to predict the state of the body and the external environment. Here, it is seen as a decision process based on the available information. The agency block (green) combines a prospective or prior component which relies on information on the action goal, and a retrospective or posterior component that reflects the discrepancy between actual and predicted sensory afference after movement completion. The Optimal controller block (blue) generates the motor commands, which reflect the task requirements, and the current state as predicted by the state observer.

is true. Specification of the action goals and task requirements and the subjective cost and benefits of an action contribute to the ‘prospective’ component of the sense of agency (Chambon et al., 2014; Moore and Haggard, 2008). As the action is initiated, the mismatch between observed and expected sensory outcomes (Blakemore et al., 2002; Frith et al., 2000) combined with the reliability of the sensory afference (Moore and Fletcher, 2012) continuously modifies such belief. In other words, the sense of agency is a dynamic process, which involves a continuing evaluation of one’s awareness of being in control. At the end of the movement, our overall (‘retrospective’) sense of agency results from the integration of these multiple sources – see Figure 2.1. In probabilistic terms, the prospective and retrospective components of the sense of agency while carrying out an action can be seen as prior and posterior probabilities of being in control. In more formal terms, prospective self agency can be seen as the subjective belief (probability) of being in control:

$$\text{prospective agency(self)} \triangleq \Pr(\text{self}) \tag{2.1}$$

before action onset. Similarly, retrospective self agency can be seen as the probability of being in control after the action is made, conditioned to all evidence collected during action (i.e. the sensory observations, i.e. the sensory inputs observed over the whole action, between time = 0 and time = T. We denote sensory observations as $y(0:T)$:

$$\text{retrospective agency(self)} \triangleq \text{Pr(self} \mid \text{y(0 : T))} \tag{2.2}$$

Bayes' theorem states that posterior and prior probability are related through the 'likelihood' of the prediction, which reflects how likely those specific sensory observations are if we assume that we are actually in control. This is quantified as the probability density function of the observation, under the 'self' assumption:

$$\text{likelihood(self)} \triangleq \text{p(y(0 : T)} \mid \text{self)} \tag{2.3}$$

Likelihood naturally emerges from the inference process underlying state estimation, but is not simply reflected in the mismatch between observed and predicted consequences of action, but more generally in the likelihood of that prediction – a quantity which, besides the mismatch, also reflects the reliability of observations – see Supplementary Notes in section 2.5. Overall, these quantities are related (Bayes theorem) as:

$$\text{retrospective agency(self)} = \frac{\text{likelihood(self)} \times \text{prospective agency(self)}}{\text{p(y(0:T))}} \tag{2.4}$$

It should be noted that the likelihood accumulates over time as the action proceeds, which suggests that the sense of agency is also a dynamic process. At movement start, the sense of agency is coincident with prospective agency before the action takes place. While the action proceeds, the sense of agency may be at times weaker or stronger. This formulation naturally extends the comparator model, and in particular implies that the sense of agency is not only strengthened by a greater match, but also by a greater reliability of the sensory information; see also (Izawa et al., 2016; Moore and Fletcher, 2012). Figure 2.1 (green block) depicts the relation between likelihood, retrospective and prospective agency.

2.2.2 From self agency to joint agency

The above framework can be extended to the sense of joint agency, intended as the subjective sense of shared control over a joint action. We can safely assume that all players are equipped with their own state observer, movement controller – hence the ability to generate their own action - and their subjective sense of self agency. In joint action each player has his/her own goals. The goals may be the same for all players; may be different, or even opposite. However, joint action implies that there is some incentive to act together, which may be a specific task requirement, e.g. the need to reach a shared goal at the same time - or is implied by mechanical coupling e.g. when dancing the waltz. In all cases, acting together always requires that each player accounts for their partner when selecting their action. Ganesh et al. (2014) examined a scenario in which two players are mechanically connected through an elastic force (a virtual spring) and are instructed to track the same moving target. The players share the same goal and, as they are physically coupled, goal fulfillment depends on both self and partner state. Chackochan and Sanguineti (2019) also focused on a scenario in which two players are mechanically connected. Both were instructed to perform reaching movements through different via-points, while at the same time keeping the interaction forces low. Again, there is physical interaction and, although the goals are distinct, they still require players' coordination. Bolt et al. (2016) focused on a situation in which two players emit sounds by pressing keys, in alternation according to different rhythmic patterns. In this case there is no physical interaction, but the players need to account for each other in order to synchronize. All the above examples suggest that in joint action scenarios the control policy must necessarily account for both self and partner's state.

The partner model Computational accounts of joint action are relatively rare, and even less (Chackochan and Sanguineti, 2019; Li et al., 2019; Pesquita et al., 2018; Takagi et al., 2018) explicitly address the need to account for both self and partner state. If there is no mechanical coupling, this can be achieved with distinct 'self' and 'other' state observers; see for instance (Pesquita et al., 2018). However, if two players interact physically, i.e. they exchange forces, their body states cannot be estimated independently. Consequently, to provide reliable estimates the 'self' and 'other' state observers must account for each other. In the words of Noy et al. (2017), 'coupled forward models are necessary for producing co-confident motion'. In conclusion, in joint action we must posit an additional module as part of the state observer, which accounts for the partner's state and possibly motor actions, which we will refer as the 'partner model' – see Figs. 3 and 4. Estimating the partner state is no different from estimating their own state, in the sense that it requires a forward model of partner's body and the availability of sensory signals – e.g. vision, proprioception, audition, etc. - which provide information about the partner. From a computational perspective, partner models may take various forms. They may just estimate the ongoing partner movements by combining measures and prior assumptions. More accurate models may additionally account for the partner's body dynamics, e.g. inertial properties. For instance they may use internal representations of the partner's body dynamics to infer their underlying motor commands (Chackochan and Sanguineti, 2019; Gillijns and De Moor, 2007). Further, partner models may be capable of inferring the partner's control policy, i.e. the mapping between the state of both players and the partner's motor command (Li et al., 2019). This type of partner model is much more informative, as it provides not only an estimate of past action, but more generally on the partner's ongoing strategy, which can be seen as a representation of the partner's ultimate goal or intentions. There is no agreement in the literature about the type of partner models and on how they are formed. During joint action players may simply estimate the ongoing partner actions (Chackochan and Sanguineti, 2019). Other studies suggest that players may develop more general partner representations, also accounting for intentions and ultimate goals (Sebanz et al., 2005). Humans are indeed very good at extrapolating higher order information by observing the motion of their peers: not only they can infer intentions or goals, but can also discriminate between actual human movements from movements that have artificial spatio-temporal features, e.g. they are incompatible with human body biomechanics (Zunino et al., 2020). Our main focus here is on 'the mutual give and take between two or more individuals involved in social interactions' (Frith and Frith, 2008), but the 'partner model' notion implies, more generally, an ability to infer the intentions, desires and beliefs of others – the Theory of Mind. We are not explicitly addressing the nature of these processes here. We posit that sensory observations are compared with their predictions under a specific hypothesis to form a likelihood. Likelihoods of the different hypotheses are combined with prior knowledge (priors) into posterior probabilities, which mediate the subsequent action generation. These computational modules constitute a basic form of Theory of Mind. The current partner model formulation focuses on implicit and immediate perceptual processes (Gallagher, 2008). More explicit and aware forms of reasoning about the partner (Carruthers and Smith, 1996; Gallese and Goldman, 1998) may contribute to higher-level prediction of partner actions, which includes intentions (i.e. the control policies) and ultimate goals (i.e. their objective function). Although these mechanisms are not covered in the current description of the model, they are clearly consistent with the overall architecture. Overall, the proposed model points at a Theory of Mind which posits fast perceptual processes and slower, more cognitively demanding processes which contribute to an efficient understanding of the partner's mental states (Frith and Frith, 2008; Gangopadhyay and Miyahara, 2015; Meinhardt-Injac et al., 2018) . In conclusion, joint action requires that the state observer has an additional 'partner model', which may be either distinct or interrelated to the 'self' state

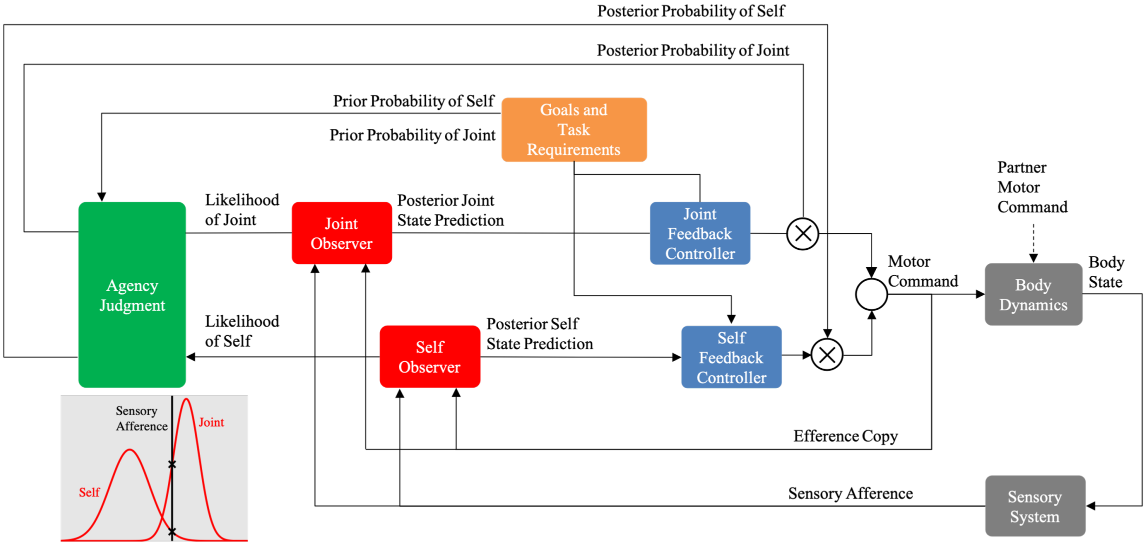

Figure 2.2 Computational model of joint agency. The assumption that agency is a decision process implies the choice between a set of options. The model involves two optimal estimator (state observer) and optimal controller pairs running in parallel. Each state observer estimates the likelihood of, respectively, self and joint. These quantities are combined with prior information from the task representations to get posterior probabilities of self and joint. These quantities gate the outputs of the optimal controllers and therefore affect the control policy.

observer – see Fig. 2.2 and Fig. 2.3. We suggest that the partner model plays a pivotal role in developing and modulating the sense of agency in joint action.

Game theory in joint action Game theory provides the analytic and computational substrate for the decision-making and control processes underlying joint action. A ‘game’ is a situation ‘involving two or more individuals whose interest are neither completely opposed, nor completely coincident’ (Nash, 1953). Concepts from game theory have been widely applied in the social sciences to understand how multiple agents coordinate their actions. In a joint action, individuals still aim at maximizing their own subjective utility, but the latter also depends on their partner’s state. This introduces additional complexity in the individual action selection mechanisms, and likely plays a role in forming our sense of joint agency. In some joint action scenarios, agents agree on a common strategy – for instance, through verbal communication - before the action takes place, thus behaving as a collective (Bacharach, 1999; Newton, 2018; Tuomela, 2007). These situations are referred as ‘cooperative games’. We argue that in these situations the sense of joint agency has a mainly ‘prospective’ character as it largely develops before action initiation. In other scenarios there is no explicit

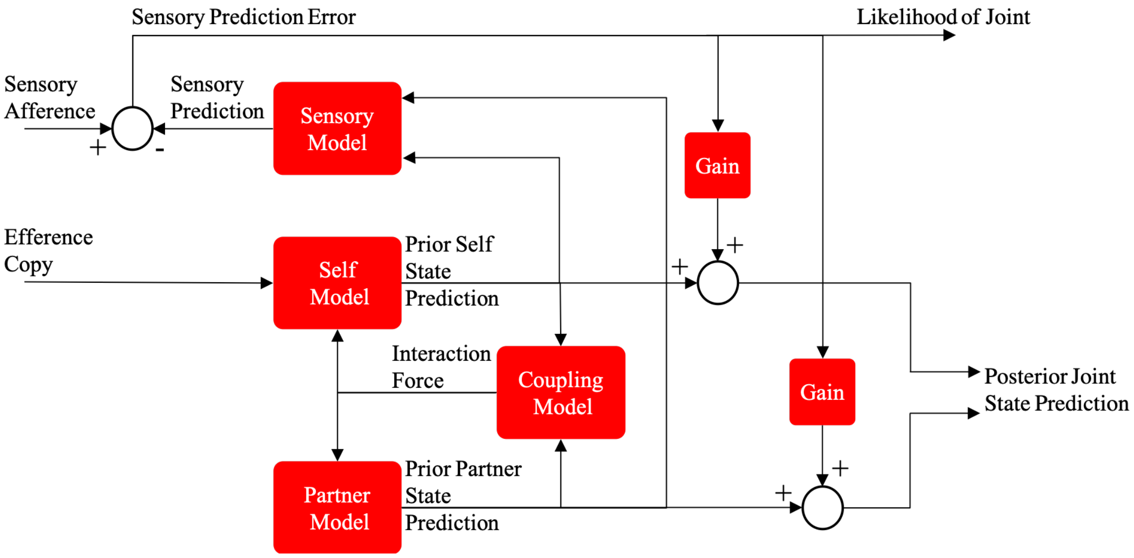

Figure 2.3 Joint state observer. In addition to a model of own body dynamics (self model), it also maintains a representation of the partner and of the self-partner mechanical coupling (if any). The observer predicts the state of both self and other. These quantities are combined to predict the sensory outcomes which are compared with actual sensory afferences. The sensory prediction error quantifies the joint likelihood. Combined with the prior, it generates the joint sense of agency.

prior agreement on a common strategy. While we know our goals, we may not be equally aware of the goals of our counterpart. In these situations, coordination emerges gradually as each agent collects information about their partner actions, their outcomes and possibly their ultimate goals, using various mechanisms (Vesper et al., 2017), during continuous or repeated interaction. Situations in which each player autonomously decides the action to take are referred to as ‘non-cooperative games’. Situations in which none of the players can unilaterally improve their benefit are known as Nash equilibria (Nash, 1951). Differential game theory is the natural extension of optimal control to joint action scenarios as it addresses those situations in which actions develop in time (Başar and Olsder, 1998). In motor control scenarios, differential non-cooperative games have been used to model situations in which humans deal with their counterpart without speaking and by communicating just through sensory cues (visual, acoustic or haptic), but they independently determine their actions (Braun et al., 2009; Chackochan and Sanguineti, 2019; Jarrassé et al., 2012; Li et al., 2019). In sensorimotor versions of the prisoner dilemma and of the rope pulling games, Braun et al. (2009) found that the players converge to Nash equilibria. In contrast, when the same task is performed by one single agent using two hands, they converged to a cooperative solution. These results suggest that, if they have perfect information about their partners, two

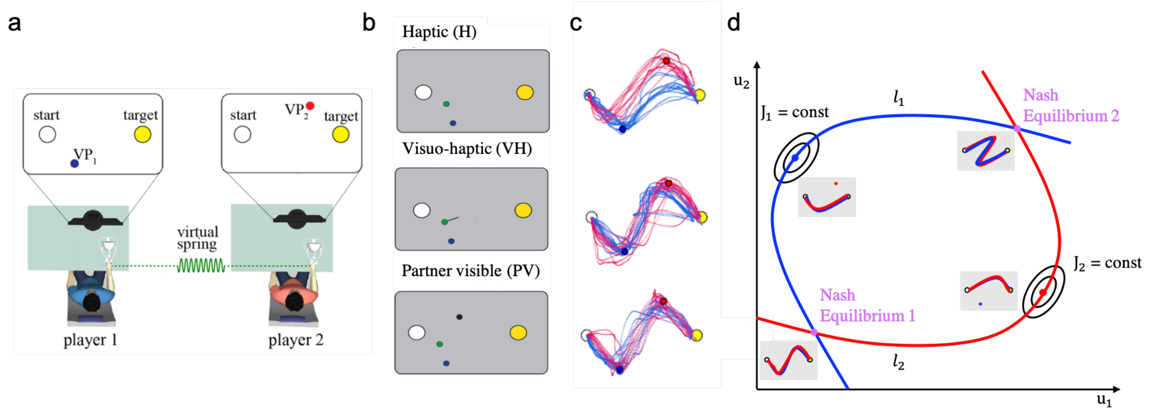

Figure 2.4 A joint action scenario and its game theoretic interpretation. a) Experimental apparatus and task. Two players are connected via a virtual spring. They were instructed to perform planar reaching movement from the same initial point to a final target, by crossing different via-points. b) Experimental conditions. Three experimental groups different in the amount of information provided about their partner: in the haptic condition (H) players could only feel their partner haptically, through the interaction force; in the visuo-haptic condition (VH), the interaction force vector was displayed on the computer screen; in the partner visible condition (PV), players could see their partner's ongoing movements through a second cursor displayed on the screen. c) Experimental results. Last ten movements of the two players after a training session. d) Game Theory predictions. Players have distinct sets of best responses to partner actions, specified by their cost function minima (J1 and J2, displayed by their contours) in the space of possible actions (player 1: u1; player 2: u2). The game has two Nash equilibria – both players passing through both via-points. The Nash equilibria are specified by the intersection of the players' reaction curves – i.e. the locus of the optimal actions which a player may take for any given action chosen by the partner. The experimental results suggest that as the available information increases (PV group), the players tend to converge to the lowest-effort Nash equilibrium. Modified from (Chackochan and Sanguineti, 2019).

players are capable of establishing stable coordination strategies which correspond to Nash equilibria. When combined with state and partner estimation, differential game theory can be used to study situations characterized by partial information (Harsanyi, 1967) in which players form conjectures about the other player's actions. In situations where the players have competing goals and need to negotiate a joint strategy, knowledge about the partner is a major determinant of the resulting coordination strategy. In a recent work in which dyads performed a sensorimotor coordination game, Chackochan and Sanguineti (2019) manipulated the information about the partner available to each player, ranging from relatively unreliable haptic information to highly reliable visual information; see Fig. 4.5. Depending on the available sensory information, they found that the players converged to different coordination strategies. When the information was more reliable, they converged to a Nash equilibrium. When the information uncertainty increased, they converged to a strategy which assumed there were no partner. Another aspect of coordination strategies, which is highly relevant to joint agency, is that they often develop gradually. Fictitious play (Berger, 2007; Brown, 1951) has been proposed as a general mechanism for the development of a stable coordination (Chackochan and Sanguineti, 2019; Grau-Moya et al., 2013). In repeated trials, subjects gradually refine their partner model and determine their best response given that partner model. This can be interpreted in Bayesian terms – at each trial, the players combine prior beliefs with the available information on the action outcome to improve their next moves. Nash equilibria are absorbing states in the fictitious play learning process (Fudenberg and Levine, 1998) - once you get there, you stay there. The sense of joint agency likely evolves as the coordination develops. For instance, the sense of joint agency may be reduced when coordination is unstable and players have difficulty in synchronizing their actions, but may increase gradually as they converge toward a stable coordination.

A modular control architecture to reconcile self and joint agency The above considerations suggest an extension of the comparator model – or rather the probabilistic model - to account for the sense of joint agency. Like self agency, the subjective experience of acting as a group would have a prospective component. The latter can be intended as the subjective evaluation of costs and benefits associated to the development of a coordination, before the latter takes place. Contextual conditions (Lafleur et al., 2020) and personal traits directly affect the subjective evaluation of tasks and goals (Vallacher and Wegner, 1989). Therefore, they both contribute to shaping the sense of agency. In probabilistic terms this can be captured by a joint agency 'prior':

$$\text{prospective agency(joint)} \triangleq \text{Pr(joint)} \tag{2.5}$$

After action completion, a 'joint' observer, involving a self and a partner model, see Fig. 2.3, evaluates the 'joint' likelihood, i.e. the extent to which the presence of multiple players explains the sensory information $y(0:T)$:

$$\text{likelihood(joint)} \triangleq \text{p(y(0:T) | joint)} \tag{2.6}$$

Finally, the retrospective sense of joint agency results from the combination of prior assumptions (prospective agency) and the match between actual and expected observations (likelihood) – see 2.5 for details:

$$\text{retrospective agency (joint)} = \frac{\text{likelihood(joint)} \times \text{prospective agency(joint)}}{\text{p(y(0:T)}} \tag{2.7}$$

How are the sense of self and joint agency related? In a joint action context, either the self or the group may be perceived as being in control of an action. As mentioned in the Introduction, empirical findings suggest that we maintain distinct senses of self and joint agency. To explain these observations, we argue that during joint action, at least two state observers are active at the same time, respectively accounting for the ‘self’ and ‘joint’ scenario. The ‘self’ observer assumes that the player is alone in contributing to the body and environment state. The ‘joint’ observer assumes that one or more partners are involved in the action. A general structure for the ‘joint’ observer is depicted in Fig. 2.3. The joint observer extends the ‘self’ observer of Fig. 2.1 in that it also includes an internal representation of partner dynamics and of their mechanical coupling (if any); see the 2.5 for details. The state predictions of self and other are then combined to form a prediction of the sensory outcome. The sensory prediction error is used to quantify the ‘joint’ likelihood. In the general – see Fig. 2.3 - the ‘self’ and ‘joint’ observers simultaneously estimate the corresponding likelihoods from the same sensory information. Combined with prior beliefs (prospective component of self and joint agency), they generate separate retrospective self and joint agency beliefs (self and joint posteriors). We also argue that the subjective awareness of being in control alone or with other partners also has an influence on action selection: we may choose different actions if we believe we are acting alone, or another player is contributing to our action. To account for this, the model posits two controllers: the ‘joint feedback controller’, which determines a motor command under the assumption of coordinating with a partner, and the ‘self feedback controller’ which assumes that the player acts alone. Each controller generates a motor command based on the state predictions made by the corresponding observer. The self and joint posterior probabilities serve as gating signals to select which motor command will be eventually generated. The overall model is summarized in Fig. 2.2. The ‘task’ is represented in terms of goals and requirements, including the requirement of coordinating with one or more partners, and possibly prior agreements with them, and can be captured by a cost function which keeps the interaction into account. This is the prior belief that a player has before the action starts – prospective components of agency. Based on the assumptions of either acting alone or jointly, the two state observers generate different state predictions on the basis of the same sensory information. An Agency Judgement subsystem combines the prior beliefs with the evidence collected during action (self and joint likelihoods), thus providing an ongoing posterior belief of either acting alone or as a group. During movements the agency judgement continuously gates the motor command generated by the ‘self’ and ‘joint’ controllers. The proposed architecture is inspired by the ‘MOdular Selection And Identification for Control’ (MOSAIC) model (Haruno et al., 2001; Wolpert and Kawato, 1998). Originally proposed to address the modularity of the motor system, the MOSAIC model posits that the motor system uses multiple observer-controller pairs, each dedicated to a certain aspect of the action and characterized by different assumptions on our body and the environment in which the action takes place. This architecture is more parsimonious than using a single enormously complex model of the external world and accounts for experimental findings suggesting that humans are capable of learning multiple internal representations, switching between them, or combining them, based on the context in which the action is occurring to determine an appropriate motor command – see Wolpert and Kawato (1998) for more details. The model relies on specialized predictors (observers) for each of the hypotheses under consideration. In joint action these hypotheses include acting alone (‘self’), just observing someone else’s action (‘other’), or acting jointly (‘joint’). The ‘self’ and ‘other’ observers require forward models of, respectively, own and other dynamics. A ‘joint’ observer requires both (plus a model of their mechanical coupling if any). The model keeps duplication of resources at a minimum as some of these modules may be shared at the implementation level. The proposed model posits a bidirectional relationship between action control and the sense of agency. In this view, agency is a dynamic decision process, which continuously evaluates the assumptions underlying each observer pair. Players reason about themselves – their goals, their role in the action, the perceived features of their body and environment, the presence of additional partners or opponents. Our behavior changes depending on whether we feel to have a joint control over an event, or we feel we are acting individually. This suggests that the sense of agency is not simply the outcome of an inference process, but also plays a role in determining the joint control policy. The proposed model implies that the sense of self agency and the sense of joint agency are distinct, though not independent. Rather, they reflect parallel mechanisms that make sense of the available information in different ways. The Agency Judgment block uses all the available information to evaluate what is more likely. One key model prediction is that the likelihood of acting jointly is always greater or equal than the likelihood of acting individually. This is a consequence of the fact that if you add a component in a model, its likelihood with reference to a specific set of observations can only be greater or equal, but never be worse than the model without that component. If there is no partner or the partner plays no role in the interaction, the ‘joint’ observer would provide no better explanation of the sensory observations than an observer which neglects the partner, hence the 'self' and 'joint' likelihoods would be the same. Conversely, if the partner plays any role in the action, this is reflected in the action's sensory consequences. As such, the 'joint' observer would provide a more likely explanation of the sensory observations than its 'self' counterpart. Hence the 'joint' observer, which also includes the partner model, always provides more accurate predictions than the 'self' observer if a partner is actually present. Conversely, if no partner were present, the two state observers would lead to identical predictions and consequently identical likelihoods. The relationship between self and joint-likelihood may explain the different experiences of joint agency in different interaction scenarios. In 'we-agency' scenarios like marching soldiers, the partner counts a lot. Therefore, the joint-likelihood is much greater than the self-likelihood. In shared-agency scenarios the two players have complementary roles and tend to be less interdependent. Therefore, joint and self-likelihoods tend to be more similar. It should be noted, however, that the overall sense of agency also depends on the priors – the prospective components of agency – which can make the 'self' posterior probability greater than its 'joint' counterpart. This occurs for instance in extreme situations when the two players do not interact at all and therefore no joint action takes place. Here we only consider the 'self' and 'joint' scenarios because they are the most relevant for self and joint agency, but the model may even include a 'other' scenario (in which the 'self' does not participate in the action). Indeed, the observers associated to each scenario, the likelihoods they generate and their combinations with the priors may be seen as a Bayesian classifier which, at any given time instant and cumulatively at the end of an action, compares posterior probabilities which can be interpreted as the sense of self and joint agency. This model can account for the attribution problems discussed in Frith et al (Frith et al., 2000).

2.3 Joint agency in a sensorimotor interactive task

To understand the implications of the proposed model, we simulated the task and the experimental protocol used in a published study (Chackochan and Sanguineti, 2019), which specifically addressed the way coordination strategies are affected by uncertainty about the partner; see Fig. 4.5 for a summary of the experimental protocol and main results. Simulations have a twofold purpose: (i) to test the internal coherence and completeness of the proposed model; and (ii) to explore model predictions as regards the sense of self and joint agency, in different phases of the experiment and different experimental conditions.

2.3.1 Task

In the actual experiment, two players were instructed to perform planar arm movements with the same start and end points, but different via-points, while mechanically coupled through a virtual spring. In particular, each player operated the handle of a robot manipulandum, whose movements were mapped into the motion of a cursor on a computer screen placed in front of each player. To simulate the virtual spring, the forces generated by the robots at the level of each end-effector were set to be proportional to the distance between the two hand positions. The participants could not see each other and could not speak. They were not explicitly informed about performing a joint task, but were instructed to keep their interaction forces low, which is an incentive to aim at coordination. The experiment involved three groups of dyads, which differed for the amount of information available about their partner. In the haptic condition (H), the players only perceived their partner through the interaction force. In the visuo-haptic condition (VH) the interaction force was also displayed on the screen, as an arrow attached to their hand cursor. In the partner-visible condition (PV), in addition to the haptic force each player could see their partner as a second cursor moving on the screen. The experimental protocol involved a total of 156 trials. The players were initially not connected (baseline phase, 12 trials). Then the virtual spring was turned on (training phase, 120 trials). Then the connection was removed again (after effect phase, 24 trials). The peculiarity of this task is that the players have different sub-goals and need to negotiate a joint strategy, but have no other cue about their partner's intentions or ongoing actions than the interaction force alone (H group) and additional visual feedback in various forms (VH, PV groups).

2.3.2 Model Implementation

We simulated the two-loop model of Fig. 2.2. The task is completely specified by a pair of quadratic cost functions, one for each player. The cost functions have four terms, respectively accounting for (i) reaching the final target; (ii) keeping the interaction force low (only present during the training phase); (iii) passing through the via-point; and (iv) keeping the mechanical effort low; see the Supplementary Notes for details - 2.5. We modeled the players as a single linear dynamical system (with Gaussian process noise) and two separate sensory systems (with Gaussian measurement noise). The H and VH groups only differed in the uncertainty (measurement noise) of the measured interaction force (lower in the VH group, in which the force is visually displayed to the players). In the PV group, the sensory system had an additional term – visual information about partner's position and velocity which further reduces the uncertainty about the partner movements. For each player, we posited two parallel observer-controller loops. The 'self' observer assumes that there is no partner. The 'joint' observer predicts the body state by additionally accounting for a partner model, which continuously estimates the ongoing partner actions. Both state observers optimally combine prior state predictions with the current sensory afferences to estimate the next state. The 'joint' observer included a partner model which additionally predicts the partner's ongoing motor command. At each time instant, the two observers compute cumulative self and joint likelihoods, which quantify whether the sensory afferences up to that time instant are best predicted by self or joint action. At each time instant, these measures are combined with a prior probability term which reflects the prior knowledge about the task (prospective sense of agency), to form dynamic self and joint agency beliefs (posterior probabilities). We assumed that all players initially have little or no cues about the presence of a partner. Therefore, in all phases and in all simulations, we assumed that the prospective sense of self agency is slightly greater than the corresponding joint agency. In particular, we set $prospective(self) = 0.6$ and $prospective(joint) = 0.4$. This choice is clearly questionable and the prospective component of agency within this task would deserve a study on its own, as personal traits (Vallacher et al., 1989) and external context (Lafleur et al., 2020) likely affect the prospective component of agency. The 'self' and 'joint' controllers have an identical structure, resulting from the same cost function. The 'joint' controller additionally incorporates the partner motor command estimated on the previous trial, whereas the 'self' controller assumes that that there is no partner (the partner motor command is identically zero). This implements the fictitious play learning model. At every time instant, both controllers generate motor commands based on the corresponding state observers. The overall motor command is calculated as the sum of the self and joint motor commands, weighed by the corresponding self and joint agency beliefs. All details of the model implementation, the values of the model parameters and the technical details of the simulation are reported in Section 2.5.

2.3.3 Simulation results

We simulated the evolution of the behavior of three dyads, respectively in the haptic (H), visuo-haptic (VH) and partner visible (PV) group, over the entire three-phase (baseline, training, aftereffect) 156-trial experimental protocol.

Effect of training and partner information We initially focused on the self and joint-likelihood (and the corresponding self and joint agency beliefs) calculated at the end of each simulated movement (retrospective agency). The simulation results indicate that during the baseline phase, when the two players are disconnected, the partner model does not contribute significantly to state estimation, hence the self and joint-likelihood are the same. Consequently, the posterior probabilities and thus the retrospective senses of self and joint agency are dominated by the task requirements, encoded by the agency priors. Therefore, retrospective self agency prevails over joint agency. During the training phase, when the two players are physically connected, the partner model significantly improves the prediction of the sensory outcomes and therefore joint-likelihood is always greater than the self-likelihood. Consequently, the joint posterior probability is greater than its self counterpart, and the joint control action is weighed much more. In conclusion, interaction with the partner is quickly incorporated in the joint observer and - through fictitious play - in the joint controller. In simulations, we assumed that both players already have a ‘connected’ forward model, i.e. they know already about the connected dynamics. This may be unrealistic, hence in actual experiments the convergence may not be that fast. Both self and joint-likelihood are affected by uncertainty about the partner. When more reliable information is available to a player – e.g. from the H to the PV condition – both joint and self-likelihood increase. The prediction of partner actions also becomes more reliable; therefore, the joint-likelihood tends to overcome the self-likelihood. As a consequence, the players increasingly perceive themselves as acting jointly; see Fig. 2.3. In conclusion, simulations demonstrate two key model predictions. First, the relative strengths of the sense of self and joint agency are determined by a combination of the task requirements and the match between predicted and observed sensory information. Second, when the information is more reliable the partner model plays a more important role in predicting the sensory outcomes, which strengthens the sense of joint agency with respect to self agency.

Self and joint agency within a trial The proposed model also predicts that both the senses of self and joint agency are time-varying within a single movement. In simulations, we additionally looked at this temporal dynamic. In all three experimental groups we took the last trial within the training phase, when the players have developed a stable joint strategy. These simulation results are summarized in Fig. 2.5. The instantaneous joint-likelihood varies little throughout the trial. As expected, it increases when more reliable sensory information is available. This confirms the prediction that reliable sensory information is a major determinant of joint agency. In contrast, self-likelihood is always smaller and changes within the movement. The discrepancy between self and joint-likelihood reflects the importance of the partner action in estimating the dyad state in different portions of the movement. A large discrepancy implies a stronger partner contribution. In the simulations, the self-likelihood of both players exhibits a minimum around 300 ms before the time

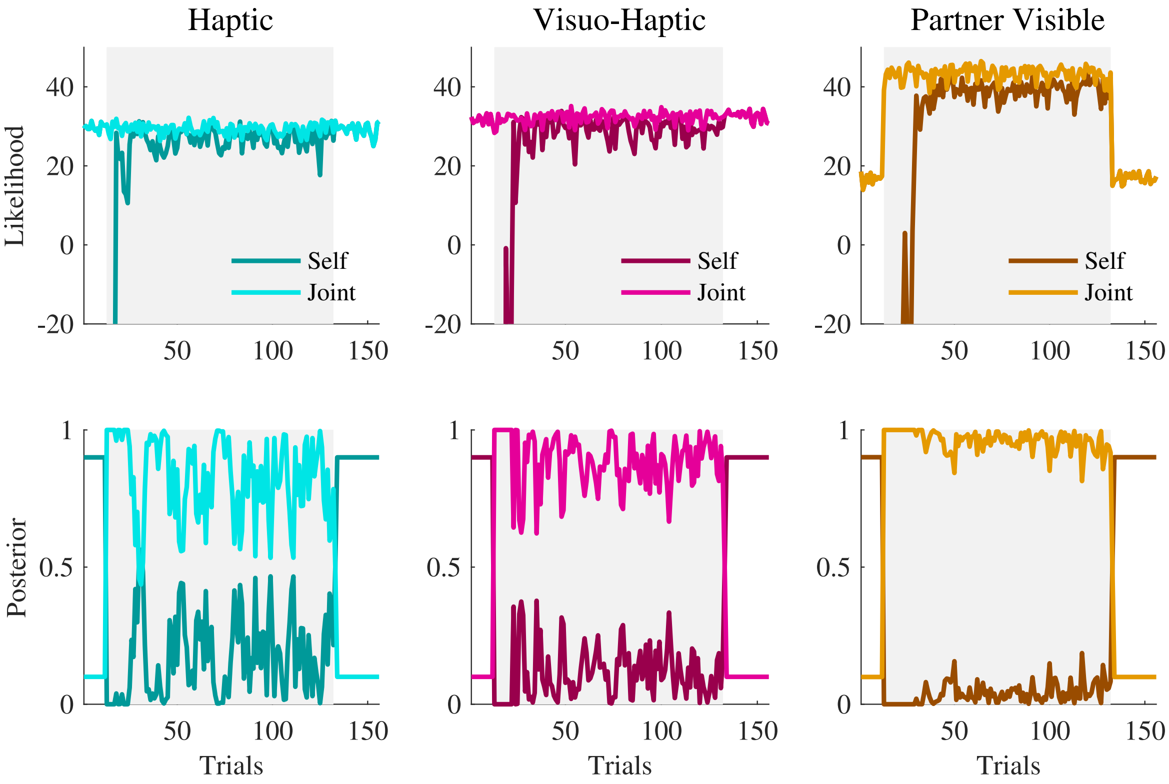

Figure 2.5 Between-trials evolution of self- and joint-likelihoods (top) and self- and joint posterior probabilities (bottom), with increasing amounts (left to right) of available sensory information: haptic (H, left), visuo-haptic (VH, middle) and partner visible (PV, right).

the partner is crossing their own via-point. Interestingly, these results are consistent with previously reported experimental findings. In particular, Chackochan and Sanguineti (2019) observed that when crossing their own via-point, each player tends to pull their partner – thus behaving like a leader, whereas their partner acts more like a follower. In our simulations, when the amount of information is greater (e.g. the PV condition), there is a greater gap between self and joint-likelihood, which points at a we-type joint agency (the players tend to lose their sense of self agency as joint agency gets stronger). In contrast, when less information is available (e.g. the H condition), both self- and joint-likelihood remain large throughout the whole trial, thus suggesting a joint agency of shared-type. In conclusion, our simulations predict that when more information about the partner is available, the joint strategy shifts from shared-type to we-type joint agency.

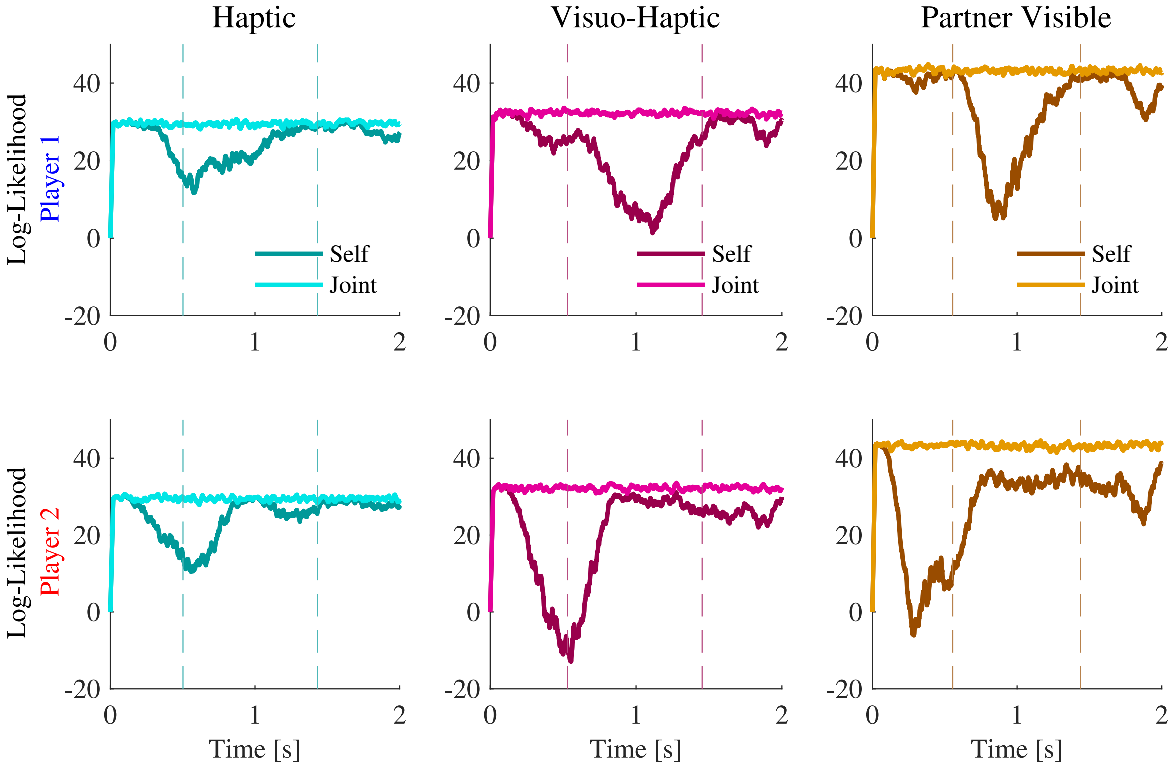

Figure 2.6 Within trial simulation results. Self- and joint-likelihood of both players within a trial (last trial in the training phase). From left to right: haptic (H), visuo-haptic (VH) and partner visible (PV) groups, for, respectively, Player 1 (top) and Player 2 (bottom). The vertical dashed lines represent the times at which each player crossed their via-point.

2.4 Discussion

We formulated a comprehensive computational framework to account for the sense of agency in joint action. By extending previous accounts (Fletcher and Frith, 2009; Izawa et al., 2016) of self agency, we suggest that both the senses of self and joint agency can be interpreted in probabilistic terms. In particular, the prospective (Chambon et al., 2014) and retrospective (Blakemore et al., 2002; Frith et al., 2000; Haggard, 2005) components of the sense of agency can be seen as prior and posterior subjective probabilities (beliefs) of being in control. Using simple arguments from the optimization framework, widely used in sensorimotor control (Todorov and Jordan, 2002; Wolpert et al., 1995), we first introduce a more general formulation of the comparator model (Blakemore et al., 2002; Frith et al., 2000) for the sense of self agency. We then argue that in order to address joint action scenarios, the state observer must explicitly account for the partner's ongoing activity (partner model) (Chackochan and Sanguineti, 2019; Li et al., 2019; Pesquita et al., 2018; Takagi et al., 2017). The partner model is essential for the development of a joint action and a sense of joint agency. It may take different forms and may have different degrees of reliability. Both aspects are central to the development of the sense of joint agency. We propose differential game theory and non-cooperative games as a general modeling framework to study joint action scenarios which develop in space and time (e.g. involving movements) and are characterized by incomplete information about the partner (Braun et al., 2009; Chackochan and Sanguineti, 2019; Jarrassé et al., 2012; Li et al., 2019). We propose fictitious play as a general learning mechanism to describe the gradual development of a stable coordination (Chackochan and Sanguineti, 2019; Grau-Moya et al., 2013) and the sense of joint agency. The proposed model is not just a mere mechanistic solution to a control problem. Its level of description is computational (what has to be computed and why) and relies on the well-established notion that the nervous system compares a variety of hypotheses in order to understand the external environment and to take decisions (Gallivan et al., 2017; Heald et al., 2018). Consistent with reports (Bolt and Loehr, 2017; Dewey et al., 2014; Le Bars et al., 2020; Pacherie, 2012) that we maintain distinct senses of self and joint agency, we posited separate observer-controller pairs, one assuming that we are acting alone ('self' loop) and another assuming that we are acting jointly ('joint' loop). The loops run continuously and in parallel. The two observers provide continuously updated subjective assessments of the 'self' (I am acting alone) and 'joint' (I am acting jointly with someone else) hypotheses, which in turn determine action selection. Each observer provides a specific interpretation. The 'agency' signal (posterior probability of either self or joint) is used to decide which controller is appropriate for the current scenario (Haruno et al., 2001). The model can be easily extended to address more scenarios, in which we just observe someone else's action ('other' scenario) or interaction with multiple agents, where the partner model combines the contributions of multiple partners (Takagi et al., 2019).

2.4.1 Model prediction

To clarify how the model works and to explore less intuitive model predictions and implications, we simulated a joint action experiment from a published study Chackochan and Sanguineti (2019).

Self and joint agency are affected by the quality of sensory information The extended probabilistic formulation of the comparator model (Frith, 2005; Frith et al., 2000; Pacherie, 2008)predicts that the amount of information available about the dyad (player and partner) modulates our sense of self and joint agency. This prediction is demonstrated by the model simulations: when sensory uncertainty is reduced, the players' ability to predict their partners' actions improves, and they feel more strongly to be part of a group, thus leading to a stronger sense of joint agency. This is consistent with reports that humans feel a stronger sense of joint agency when they coordinate their actions with a more predictable partner (Bolt and Loehr, 2017).

Are there different types of joint agency? The proposed model posits that we experience distinct senses of self and joint agency, but they are not independent. Pacherie (2012) argued that there are two qualitatively different types of sense of joint agency. In 'we-agency' (Gallotti and Frith, 2013; Pacherie, 2012), strong sense of joint agency is associated to a weak sense of self agency. In 'shared-agency' a strong sense of joint agency is compatible with a strong sense of self agency (Loehr, 2018; Pacherie, 2012). Our proposed model captures both scenarios, with no need to assume different 'types' or 'modes' of the sense of joint agency (Pacherie, 2012). We suggest that the relation between the senses of joint and self agency is determined by the task requirements and the predictability of partner's actions, with no need to assume qualitatively different 'modes' of the sense of joint agency. In particular, the 'we-agency' scenarios corresponds to situations – or phases of a joint action – in which our partner's actions have a strong impact on our sensory afferences. In this case, the joint-likelihood would be much greater than the self-likelihood, which would place a bias toward joint agency. For instance, singers in a choir with different voices coordinate their actions to achieve a shared goal and equally contribute in its fulfillment. As such they are likely to experience a greater sense of joint agency and a comparatively lower sense of self agency as each individual only contributes a fraction of the overall performance. Conversely, 'shared-agency' scenarios correspond to situations in which partner actions play a weaker role in the overall coordination, or have different impacts in different phases of the movement, e.g. when the players have different roles. In this case the partner information would not add much to the joint-likelihood with respect to the self-likelihood. Similarly, the model predicts that we experience different proportions of sense of joint and self agency in competitive situations with respect to collaborative scenarios. For instance, two fighters likely experience a lower sense of joint agency with respect to two partners lifting a heavy object together. Joint weight lifting implies a strong prospective sense of joint agency and a high joint-likelihood, hence a strong retrospective sense of joint agency as compared to self-agency. Fighting implies a strong prospective sense of joint agency, but the joint-likelihood would be low, as the competitor would aim at being less predictable. As a consequence, fighting implies a weaker retrospective sense of joint agency (as compared to self-agency). These predictions are also consistent with (Silver et al., 2021), which identify the level of cooperation as the key action feature which modulates the sense of joint agency. The model provides a computational substrate for the concepts of we-ness and we-representation (Gallotti and Frith, 2013; Kourtis et al., 2019). A partner in a joint action gathers significance only if he/she is seen in the self-perspective, as an augmentation of the self. We would not need to account for our partner if there are no requirements to undergo a joint action, as in the unconnected phases of the simulated experiment.

Agency as a dynamic experience Our proposed model suggests that both senses of joint and self agency evolve in time. Before an action takes place, an early (prospective) sense of agency is determined by the task goal, requirements and constraints, and additional cognitive and emotional factors, as suggested by Synofzik et al. (2013) and Chambon et al. (2014). This initial bias – the self and joint priors in our model – is combined with self and joint-likelihood signals, each quantifying the reliability of a different predictor of the ongoing action. The relative magnitudes of self and joint-likelihood change within a single action, up to action completion (retrospective agency). In the simulations, each player coordinates with their partner, but also needs to maintain their identity to achieve their own sub-goal – i.e. crossing their respective via-point. When the sensory information is more reliable, the sense of joint agency becomes stronger. However, the sense of self agency only increases if the partner is not perceived to effectively contribute to the action.

Agency determines the decision/control process In our proposed multi-loop control architecture, the experienced self and joint agency play a pivotal role in motor command selection and therefore in the development of the joint strategy. In other words, the sense of agency is not just about perception, but also affects action. The ‘self’ and ‘joint’ state observers give rise to separate controllers, whose relative importance is weighed by the sense of self and joint agency. Hence the sense of agency (self and joint) is not simply a byproduct of perception, but emerges from this dynamic process which involves a bidirectional interplay between action control and estimation of own and partner state. A strong sense of joint agency places more emphasis on more collaborative actions and facilitates the convergence to a stable coordination, in which each player accounts for their partner and selects their action accordingly. If information about the partner is more uncertain, the sense of joint agency is lower and the players converge to a strategy which relies less on partner’s actions. In both cases, the selected actions further emphasize either the senses of joint or self agency.

2.4.2 Learning in joint action

An open issue in joint action is how two players converge to a coordination strategy. Fictitious play assumes that the players select the action which maximizes their respective objective function, by accounting for the empirical probability distributions of the actions of their opponent. A key property of fictitious play is that it only requires that each agent knows their partner's probability distribution of previous actions. Higher-order theories of mind, possibly representing partner intentions (i.e. the control policies) or goals (i.e. their objective function) would lead to more efficient learning and/or to different Nash equilibria. Higher-order forms of theory of mind may lead to an infinite regression - each agent represents their opponent's policy which accounts for their own and so on; see also Yoshida et al. (Yoshida et al., 2008). If an agent does not know in advance the relation between action and its subjective 'value' (in terms of our model, the 'task representation' block), he/she can learn it through some biologically plausible form of reinforcement learning, like Q-learning (Claus and Boutilier, 1998). In joint action scenarios, this is more complicated as each agent's goal – and its subjective value - depends on the actions of both agents. Here a question arises of whether learning the action(s)-value and predicting the partner actions occur independently, or are actually part of one single process. In a recent study focusing on sensorimotor versions of classical discrete games, Lindig-León et al. (2021b) observed that convergence to a Nash equilibrium is consistent with a model-free form of reinforcement learning, in which actions are generated as a trade-off between their value and the requirement of minimizing their change with respect to the previous trial. This learning mechanism does not explicitly account for partner actions, but it is unclear if it would extend to more complex forms of coordination that involve more than just discrete decisions. In conclusion, how a coordination strategy is learned is an open issue, and the value of the fictitious play hypothesis is that it represents a reference hypothesis against which other possibilities can be tested empirically.

2.4.3 Relation with previous models of joint action

As noted in section 2.2.2, the proposed model is inspired to the MOSAIC computational architecture (Haruno et al., 2001; Wolpert and Kawato, 1998), which was originally developed to address motor system modularity. MOSAIC posits multiple observer-controller pairs, each specialized for specific actions (e.g. grasping a specific object) or environmental conditions. In addition to the observer and the controller, each module has a responsibility predictor which uses contextual information to estimate the prior of that module. Priors and likelihoods (provided by the observer) are then combined into a weighing signal for the motor commands calculated by each controller module. The model was later extended to address sensorimotor learning (Haruno et al., 2001) and action observation (Wolpert et al., 2003). This latter study suggested that one pair of such models – one for action control, and the other devoted to action observation - could capture the mechanisms underlying social interaction. A recent study (Haar and Donchin, 2020) discusses MOSAIC's possible neural substrates. Here we extend the MOSAIC model by (i) incorporating priors which, combined with the likelihoods, provide the posterior probability of each scenario; and (ii) allowing for loops reflecting self and joint action. Self and joint action are treated as different scenarios, and self and joint agency are determined as their respective posterior probabilities. Similar to 'responsibilities' in the MOSAIC model, the agency signals affect the action selection process. The Predictive Joint Action Model, PJAM (Pesquita et al., 2018) specifically addresses joint action. PJAM builds on HMOSAIC, an evolution of the original model (Haruno et al., 2003) which addresses the multilevel character of perception and action through a hierarchy of MOSAIC layers. Each layer receives posterior probabilities from the lower level module (bottom-up path), which specify the currently selected module in the current behavioral situation. The output of higher-level modules is a set of (top-down) prior probabilities of the subordinate modules, which act to prioritize lower-level module selection. PJAM posits a hierarchy of 'predictive processors' (i.e. observer-controller pairs), which captures multiple levels of action representation, from high-level (symbolic) to low-level (movement feature). Each layer posits a 'self' and a 'other' observer-controller loop. At higher levels in the hierarchy the loops merge into a 'joint' loop, whereas at lower levels of the hierarchy the 'self' and 'other' loops remain distinct. Although the architecture is only described qualitatively, it seems effective in capturing situations in which the incentive to act together is encoded into each player's objective function and is evaluated before the action takes place. PJAM is not intended for generating motor commands in a joint action context, but rather as a module for hierarchical predictive processing, focusing on action planning. With respect to PJAM, our proposed model focuses on the generation of motor commands aimed at gradually achieving a stable coordination with a partner, even in situations in which there is minimal or no information about partner's actions or intentions. This necessarily requires a 'joint' controller, which explicitly accounts for both the 'self' and 'other' states. We also suggest that the controller block is an optimal feedback controller, a more general formulation than MOSAIC or PJAM, which is easily extended through differential game theory to address joint action controllers. Likewise, on the prediction side we point out that in general the 'self' and 'other' state predictions cannot run independently, even at lower levels of representation – think for instance of mechanical coupling. This suggests that a general hierarchical predictor-observer architecture must involve a 'joint' loop at all levels of representation. In other words, a 'joint' predictive processor (or equivalently two inter-dependent self-other predictive processors) is required at all levels of the hierarchy. In our model, we added a 'self' observer-controller loop to the 'joint' loop to account for the empirical observation of distinct experiences of joint and self agency. In conclusion, our proposed model differs from PJAM in that it posits a 'joint' and 'self' rather than a 'self' and 'other' loop. Another difference is that our model does not address the multilevel character of action representation and only focuses on the lower level (movement feature), but the extension is straightforward. Finally, and distinctively, we suggest that the responsibility signals can be interpreted as distinct subjective 'joint' and 'self' agency beliefs, an aspect which is not specifically addressed by HMOSAIC or PJAM. In conclusion, our proposed model departs significantly from PJAM. We believe it provides a coherent picture of the inter-relation of prediction, control, and the sense of agency in a broader range of joint actions. Recently, in the context of a gaze-contingent virtual task, Brandi et al. (Brandi et al., 2019) proposed a Bayesian model for "social agency", defined as the experience of control over the social environment. Based on experimental work of Pfeiffer et al. (Pfeiffer et al., 2012), the model relies on specific key features of partner action, namely gaze direction (e.g. making eye contact) and temporal responsiveness. They propose that the (precision-weighted) mismatch between these observations and their predictions determines the degree of social agency. The precision-weighted sensory prediction error can be interpreted as the logarithm of the likelihood, hence the model is very related to the one proposed here. In fact, in a Bayesian framework the specific features that are more relevant to joint agency automatically emerge as the main contributors to the likelihood, with no need to specify them a priori. Our proposed model can be related to the free energy principle (FEP) framework (Friston et al., 2006). FEP aims at unifying perception and action in terms of the minimization of one single quantity – free energy. Like the FEP, our proposed model relies on an optimization framework, but assumes separate costs for action generation and perception. Recently, Kahl and Kopp (Kahl and Kopp, 2018) proposed a hierarchical FEP-based model of sensorimotor coordination, which is capable to discriminate between self and other actions. Self and other agency are simulated in a handwriting task, by manipulating perception so that it is either consistent or not with their own action. They found that inconsistent (i.e. other) behavior corresponds to a greater free energy magnitude. However, the model does not directly address joint action and it is unclear how it would select the correct action in 'joint' scenarios in which the players are required to act at the same time, like in continuous coordination games. All the above models (Brandi et al., 2019; Haruno et al., 2003; Kahl and Kopp, 2018; Pesquita et al., 2018) propose a hierarchical architecture to address the different levels of abstraction of the perception-action process. Here we only focus on a minimal architecture which focuses on the bottom layer of the prediction-action cycle, but the proposed framework can be extended to multiple layers to account for the integration of partner models at different levels of description.

2.5 Conclusion

Based on a Bayesian probabilistic framework we propose a general model of joint action which brings together prediction, control and the sense of agency in a broad range of joint action scenarios. The model extends previous models of the sense of self agency with two main extensions: a partner model which is capable of predict the partner actions or intentions as part of a 'joint' state observer, and an optimal controller based on differential game theory. The model posits two observer-controller loops which run in parallel, accounting for 'joint' and 'self' prediction and control and provide continuous estimates of the experienced sense of joint and self agency. The proposed probabilistic framework captures the inter-individual variability in joint action, so that the sense of agency is a highly subjective experience. Variability affects the sensorimotor loop at different levels. Individual differences are apparent in sensory acuity and in the accuracy of our internal representations, and personal traits affect the subjective evaluation of goals and task (Vallacher and Wegner, 1989), leading to a more distal or proximal experience of action. These features are accounted by the priors and by the reliability of the sensory channels. The proposed model has several speculative aspects which call for empirical test. The nature of the partner ('other') model is debated. It may only focus on ongoing movements; or the ongoing motor commands; or even the underlying intentions/goals. This aspect probably depends on the task and its context, and can be explored empirically. Another aspect is the relation of the partner model with the prediction of our own action ('self' observer). Different to similar accounts (Pesquita et al., 2018) we suggest that 'self' and 'other' observer are interdependent, so that they should be really considered as a single 'joint' observer. How is a stable coordination strategy developed is another aspect of the proposed model. We proposed fictitious play as a general mechanism, but there are alternatives – e.g. reinforcement learning – which may better capture the experimental observations and could provide additional insights about the prospective and retrospective components of the sense of agency. Another aspect of the model that call for additional empirical test is the introduction of a 'self' observer-controller loop as the basis to account for the empirical observations of distinct senses of joint and self agency: how does it compare with the accounts of we-agency and shared-agency subtypes of joint agency?

Differential game theory in conjunction with optimal estimations provides a wide range of tools which can be used to address joint action The design of experiments to disentangle the senses of self and joint agency is extremely challenging. Questionnaires provide useful tools to quantify the sense of self and joint agency, but they have limitations in capturing the temporal evolution of the agency during a sustained joint action performance and in the independent assessment of joint and self agency. Other approaches, like intentional binding (Haggard et al., 2002), need to be extended to specifically address joint agency. Overall, computational and empirical approaches can potentially inform each other for a complete understanding and a general description of the sense of agency and its behavioral implications.

Supplementary Notes

This section describes the technical details of the computer simulations which are reported in the main sections of the current chapter. The general purpose of the simulations is to characterise the evolution of probabilistic indicators of both senses of joint and self agency within a trial and across trials within a training protocol which corresponds exactly to the experiments reported by (Chackochan and Sanguineti, 2019).

Modeling framework

The modeling framework relies on the formulation reported in (Chackochan and Sanguineti, 2019). We modeled the dynamics of the bodies of two players who are physically connected by a spring and their task as a differential non-cooperative game, with Gaussian noise and quadratic costs. We modeled each player arm and robot dynamics as a point mass m_i, $i = 1,2$.

$$m\ddot{p}_i = f_i + k \cdot (p_{-i} - p_i) - b \cdot \dot{p}_i + m \cdot g \tag{2.8}$$

Where p_i and p_{-i} are players hand position, m_i is the player's mass and f_i is the muscle force. k is the elastic constant ($k = 150\, N/m$) which define the interaction force. Each player is also subject to a viscous force (viscous constant $b = 10\, Ns/m$) and to gravity g. For simulations we configure $m_1 = m_2 = 2\ Kg$. The muscle force dynamics is modeled as a second order system:

$$\tau^2 \ddot{f}_i + 2\tau \dot{f}_i + f_i = u_i \tag{2.9}$$

Where τ is the time activation constant ($\tau = 40ms$). And u_i is the activation vector, which constitutes the system input. Therefore the dyad dynamics in state-space form is:

$$x(t+1) = A \cdot x(t) + B_1 \cdot [u_1(t) + \eta_1(t)] + B_2[u_2(t) + \eta_2(t)] \tag{2.10}$$

where $x = [p_1^T, \dot{p}_1^T, f_1^T, \dot{f}_1^T, p_2^T, \dot{p}_2^T, f_2^T, \dot{f}_2^T]^T$ is the state of the body and the environment (including position and velocity of both players) and $u_1(t)$ and $u_2(t)$ are the two players' motor commands. Parameters A, B capture the players' motion equations.

$$A = \begin{bmatrix} 0_2 & I_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 \\ -k/m_1 & -b/m_1 & 1/m_1 & 0_2 & k/m_1 & 0_2 & 0_2 & 0_2 \\ 0_2 & 0_2 & 0_2 & I_2 & 0_2 & 0_2 & 0_2 & 0_2 \\ 0_2 & 0_2 & -I_2/\tau^2 & -2I_2/\tau & 0_2 & 0_2 & 0_2 & 0_2 \\ 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & I_2 & 0_2 & 0_2 \\ k/m_2 & 0_2 & 0_2 & 0_2 & -k/m_2 & -b/m_2 & 1/m_2 & 0_2 \\ 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & I_2 \\ 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & -I_2/\tau^2 & -2I_2/\tau \end{bmatrix} \tag{2.11}$$

$$\begin{bmatrix} B_1 & | \, B_2 \end{bmatrix} = \begin{bmatrix} 0_2 & 0_2 \\ 0_2 & 0_2 \\ 0_2 & 0_2 \\ I_2/\tau^2 & 0_2 \\ 0_2 & 0_2 \\ 0_2 & 0_2 \\ 0_2 & 0_2 \\ 0_2 & -I_2/\tau^2 \end{bmatrix} \tag{2.12}$$

Where $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $O_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. Each player's motor command was assumed to be affected by an additive zero-mean Gaussian noise (motor noise): $\eta_i(t) \sim N(0, \Sigma_i^{\eta})$. Each player has their sensory system which provides the sensory feedback y_i and is described as a linear shift-invariant dynamical system:

$$y_i(t) = H_i \cdot x(t) + v_i(t) \tag{2.13}$$

Where parameters H_1,H_2 encode the geometry of each player's sensory system. The size of H_i depends on information availability. For instance, in H and VH condition the sensory information $y_i = [p_i, \dot{p}_i, k \cdot (p_{-i} - p_i), x_T, x_{VP_i}]^T$ and the H_1 matrix take the following form:

$$H_1 = \begin{bmatrix} I_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 \\ 0_2 & I_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 \\ -k \cdot I_2 & 0_2 & 0_2 & 0_2 & -k \cdot I_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 \\ 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & I_2 & 0_2 & 0_2 \\ 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 & I_2 & 0_2 \end{bmatrix} \tag{2.14}$$

Instead for the PV group the sensory feedback contains also visual information about partner position and velocity $y_i = [p_i, \dot{p}_i, p_{-i}, \dot{p}_{-i}, k \cdot (p_{-i} - p_i), x_T, x_{VP_i}]^T$ and the parameter H_i changes accordingly. The noise associated to different sensory channels is assumed to be Gaussian with variance:

$$\Sigma_i^v = diag(\sigma_x^2, \sigma_x^2, \sigma_x d^2, \sigma_x d^2, \sigma_f^2, \sigma_f^2, \sigma_x^2, \sigma_x^2, \sigma_x^2, \sigma_x^2) \tag{2.15}$$

With $\sigma_x^2 = 1.7^2\ mm^2$, $\sigma_{xd}^2 = 35^2\ mm^2/s^2$. While $\sigma_f^2 = 2^2\ N^2$ for H and PV groups and $\sigma_f^2 = 0.05^2\ N^2$ for VH group.

Control Architecture

To study how the senses of self and joint agency develop within a sensorimotor interactive task (Chackochan and Sanguineti, 2019), each player has two parallel motor control loops. One assumes the player is acting alone and the other that they are acting in a joint configuration. Each loop integrates the optimal state estimation and the optimal control, with a state observer and a feedback controller. Each state observer predicts the future state which are used to compute the sense of agency based on priors and current likelihood. The estimated state is input in the respective feedback control which compute the motor command. Motor commands from the two feedback controllers – self and joint – are combined to issue a final motor command which reflects the sense of agency. Therefore, the sense of agency is computed instantaneously, is retained in the memory and affects the action performance.

State observers

We assumed that each player has two state observers. Both observers are modeled as a Kalman filter, optimally combining a prediction (prior) and a correction or innovation term:

$$\hat{X}_i^+(t) = \hat{X}_i^-(t) + W(t) \cdot [y_i(t) - \hat{y}_i(t)] \tag{2.16}$$

The innovation term is proportional to the displacement between the actual sensory signal, $y_i(t)$ and its prediction, $\hat{y}_i(t) = H_i \cdot \hat{x}^-(t)$.

[Notation: The hat symbol ^ over a variable name denotes an 'estimate'. The '-' and '+' signs indicate, respectively, prior and posterior estimates.] In the 'joint' observer the partner model may account for the partner's full body dynamics, e.g. inertial properties, thus estimating the past motor commands. In this case, the joint state observer maintains a model of the whole dyad dynamics – which includes both self and partner, as well as their mechanical coupling; see Fig. 4.5 in the main sections of the chapter. Alternatively, the partner model may simply account for the partner's movements but does not explicitly account for the body dynamics and the underlying motor commands. The joint state observer estimates the state of both self and partner. In our simulations, we assumed that the partner model fully accounts for the partner's dynamic behavior. In particular, in the 'joint' observer, the observed state $X_i = [x\ u_{-i}]^T$ combines the actual state x of the body and the partner's motor command, u_{-i}. In other words, the joint observer estimates both the state of the body and the partner's current action. Prior knowledge about the plant dynamics, an efferent copy of his/her own motor command, and prior knowledge about the partner allow to make predictions about the next plant state. Specifically, the prediction term accounts for the body dynamics, the player's own input and the partners's action.

$$\hat{x}_i^-(t) = A \cdot \hat{x}_i^+(t-1) + B_i \cdot u_i(t-1) + B_i \cdot \hat{u}_{-i}^+(t-1) \tag{2.17}$$

In the partner model, we simply assumed that the prior estimate of the partner's action is just the posterior predicted action at the previous time instant:

$$\hat{u}_{-1}^-(t) = \hat{u}_{-i}^+(t-1) \tag{2.18}$$

In the 'joint' observer, estimation of partner actions is based on the combination of such prior knowledge (a constant motor command) and the correction provided by the sensory afference. This mechanism relies on the notion that sensory information contains information about partner's previous control signal. Therefore, like the state estimate, partner' action is estimated as a combination of prediction and correction. In the 'self' observer, the state reflects the body state, i.e. $X_i = x$. The prediction term accounts for the body dynamics, the player's own input but not the partner's:

$$\hat{x}_i^-(t) = A \cdot \hat{x}_i^+(t-1) + B_i \cdot u_i(t-1) \tag{2.19}$$

Both state observers additionally calculate at each time instant the uncertainty (covariance) of the state estimate, before and after the sensory information is provided, i.e. prior and posterior: $P^-(t) = cov\{\hat{x}_i^-(t)\}$ and $P^+(t) = cov\{\hat{x}_i^+(t)\}$ and the optimal weight $W(t)$ of the innovation term through an iterative optimal estimation procedure known as the Kalman algorithm.

Optimal controllers

The optimal controllers for each player were calculated from a task representation expressed by a strochastic, quadratic cost function (one for each player, $i = 1,2$; i: player, -i: partner):

$$\begin{aligned} J_i[u_i, u_{-i}] = w_p \cdot ||x_T - x_i(T)||^2 + w_v \cdot ||\dot{x}_i(T)||^2 + w_v p \cdot ||x_{VP_i} - x_i(tc_i)||^2 + \\ w_f \cdot 1/T \cdot \sum_{t=1}^{T} ||x_{-i}(t) - x_i(t)||^2 + r \cdot w_v p \cdot 1/T \cdot \sum_{t=1}^{T} u_i(t)^2 \end{aligned} \tag{2.20}$$

The cost functions combine requirements about body state and reflects effort minimization associated to the action. In particular, the first three terms reflect the stopping at the target and the via-point crossing. The fourth term reflects the minimization of the interaction force while the last term penalizes the costs associated to the action, representing the trade-off between task-related accuracy and effort. The weighing coefficients determine the relative importance of the requirements, and are set following the Bryson's rule (Bryson, 1975). Cost functions can be rewritten in a synthetic form:

$$J_i[u_i, u_{-i}] = E \sum_{t=1}^{T-1} [x(t)^T \cdot Q(t) \cdot x(t) + u_i(t)^T \cdot R(t) \cdot u_i(t)] + x(T)^T Q(T) \cdot x(T) \tag{2.21}$$

Where$Q(t)$ contains weights for interaction forces and reaching via-point and target. While $R(t)$ contains weights of efforts associated to the action. The resulting 'self' optimal controller has a feedback (corrective) component, which is proportional to the current body state estimate:

$$u_i(t) = -L_i(t) \cdot \hat{x}(t) \tag{2.22}$$

The 'joint' optimal controller has an additional feedforward (anticipatory) component, which reflects the estimated partner action on the previous trial:

$$u_i(t) = -L_i(t) \cdot \hat{x}(t) + l(t) \tag{2.23}$$

In this way, each players gradually adjust their motor commands to their partner's estimated action on the previous trials. Learning is modeled as fictive play, where players chose their best next action, at each trial, by assuming partner's action stationary. Within this learning algorithm Nash equilibria are absorbing states (Fudenberg and Levine, 1998).

Sense of agency

The proposed model assumes that the sense of self and joint agency for player i at each given time instant is assessed as the posterior probabilities of either acting in 'self' or 'joint' mode. From Bayes theorem, these are calculated as:

$$Pr\left(\text{joint} \mid y_i\left(1\colon t\right)\right) = \frac{p\left(y_i\left(1\colon t\right) \mid joint\right) \cdot Pr\left(joint\right)}{p\left(y_i\left(1\colon t\right)\right)} \tag{2.24}$$

$$Pr\left(\text{self} \mid y_i\left(1\colon t\right)\right) = \frac{p\left(y_i\left(1\colon t\right) \mid self\right) \cdot Pr\left(self\right)}{p\left(y_i\left(1\colon t\right)\right)} \tag{2.25}$$

where $y_i\left(1\right)$ denotes the time series of the sensory afferences $y_i(t)$ since the beginning of the movement (t=1) until time t. The quantities $Pr(joint)$ and $Pr(self)$ (prior probabilities of joint and self) reflect the task requirements and constraints, i.e. whatever contributes to the sense of agency before the action begins (i.e. the 'prospective' component of agency). In the simulations, we set $Pr(self) = 0.6$ and $Pr(joint) = 0.4$ for all the simulations, as it is reasonable to assume that at the beginning of the experiments players were not aware that they were interacting. The quantities $L_{joint}(1\colon t) = p(y_i(1\colon t) \mid joint)$ and $L_{self}(1\colon t) = p(y_i(1\colon t) \mid self)$ are, respectively, the 'joint' and 'self' likelihoods. They quantify, at a given time instant, how likely the player is acting as a 'self', or is participating in a joint action. This ongoing component of sense of agency is determined not only by the sensory prediction error but also from the variability of the state estimation. These quantities can be calculated, at each time step, in terms of the instantaneous likelihoods, $L_{joint}(t) = p(y_i(t) \mid joint)$ and $L_{self}(t) = p(y_i(t) \mid self)$ from the outputs of the two state observers. In particular, the instantaneous 'joint' likelihood of partner i is computed as:

$$L_{joint}(t) = p(y_i(t) \mid joint) = (2\pi)^{-\frac{d}{2}} |\Sigma_y(t)|^{-\frac{1}{2}} e^{-\frac{1}{2}[y_i(t)-\hat{y}_i(t)]^T \cdot \Sigma_y^{-1}(t) \cdot [y_i(t)-\hat{y}_i(t)]} \tag{2.26}$$

where $\hat{y}_i(t)$ is the sensory prediction at time t, computed from the prior state estimate and the sensory system matrix: $\hat{y}_i(t) = H_i \cdot \hat{x}_i^-(t)$ and $\Sigma_y(t)$ is the covariance measure of the sensory prediction, i.e. $\Sigma_y(t) = \Sigma_i^v + H_i \cdot P_i^-(t) \cdot H_i^T$. The total likelihoods $L_{joint}(1\colon t)$ and $L_{self}(1\colon t)$, were finally calculated iteratively from the instantaneous likelihoods as:

$$L_{joint}(1:t) = L_{joint}^{\alpha}(1:t-1) \cdot L_{joint}^{(1-\alpha)}(t) \tag{2.27}$$

Parameter α is a discount factor, accounting for the relative importance of the contributions from earlier time instants. In all simulations we set $\alpha = 0.4$. The posterior probabilities of either acting in ‘self’ or ‘joint’ mode were then used to ‘weight’ the outputs of, respectively, the ‘self’ and ‘joint’ controllers. All simulations were performed using Matlab and Simulink (Mathworks).

Chapter 3

Artificial Partners to Understand Joint Action: Representing Others to Develop Effective Coordination

In the last years, artificial partners have been proposed as tools to study joint action, as they would allow to address joint behaviors in more controlled experimental conditions. Here we present an artificial partner architecture which is capable of integrating all the available information about its human counterpart and to develop efficient and natural forms of coordination. The model uses an extended state observer which combines prior information, motor commands and sensory observations to infer the partner's ongoing actions (partner model). Over trials, these estimates are gradually incorporated into action selection. Using a joint planar task in which the partners are required to perform reaching movements while mechanically coupled, we demonstrate that the artificial partner develops an internal representation of its human counterpart, whose accuracy depends on the degree of mechanical coupling and on the reliability of the sensory information. We also show that human-artificial dyads develop coordination strategies which closely resemble those observed in human-human dyads and can be interpreted as Nash equilibria. The proposed approach may provide insights for the understanding of the mechanisms underlying human-human interaction. Further, it may inform the development of novel neuro-rehabilitative solutions and more efficient human-machine interfaces.

Reprinted from: De Vicariis, C., Pusceddu, G., Chackochan, V. T., Sanguineti, V. (2022). Artificial Partners to Understand Joint Action: Representing Others to Develop Effective Coordination. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 30, 1473-1482. DOI: 10.1109/TNSRE.2022.3176378

3.1 Introduction

Joint action is pervasive in our daily life – two persons carrying a heavy object together or folding a sheet, a therapist interacting physically with a patient are just a few examples (Sebanz et al., 2006). Joint action implies a dynamic interplay between self and other (Dumas and Fairhurst, 2021); within an ensemble, individuals adapt their behaviors to adeptly exchange information and coordinate with others (Curioni et al., 2019; Vesper et al., 2017). As their behaviors are interdependent, it is difficult to experimentally disentangle the individual contributions and the way each participant reacts to their partner's actions.

A way to overcome this problem is to design experiments in which human participants interact with simulated, human-like artificial partners (AP), whose personal traits and interaction strategies can be manipulated experimentally (Dumas et al., 2014a). This approach is often referred to as Virtual Partner Interaction (VPI) (Kelso et al., 2009) or Human Dynamic Clamp (HDC)(Dumas et al., 2014a), as it resembles the dynamic clamp technique in cellular neuroscience (Sharp et al., 1993a) which uses virtual, simulated ion channels with specific biophysical properties to understand the complexity of neuron dynamics. The original HDC approach relied on a theoretical model of coordination dynamics which involves coupled non-linear oscillators – the Haken-Kelso-Bunz (HKB) model (Haken et al., 1985). Originally introduced to model the dynamics of the relative phase between two fingers or limbs performing rhythmic movements, the HKB model was later extended to capture basic aspects of social coordination between two individuals. Early applications of the HDC concept (Dumas et al., 2014a; Kelso et al., 2009) focused on real-time, rhythmic bidirectional interaction and involved human subjects coordinating their hand movements with an avatar. In conjunction with high-density EEG recordings, this paradigm has been used to investigate the neural foundations of social interaction (Dumas et al., 2020). The VPI/HDC paradigm has provided important contributions to the study of discrete joint action. However, crucial determinants of sensorimotor interaction, like perception, decision-making and control mechanisms need to be explicitly modeled to develop more versatile artificial agents. APs have been also used in applications involving decision-making, with no actual movements. For instance, when a human participant interacts with a simulated partner whose risk-sensitivity is modulated depending on model uncertainty (Grau-Moya et al., 2013).

Here we propose a biomimetic artificial partner (AP) architecture which can develop collaborative strategies with a human partner. The model builds upon previous work (Chackochan and Sanguineti, 2019) on computational modeling of joint action. In particular, we assume that the AP behaves optimally at both perceptual and control level.

These assumptions rely on a large body of evidence from perceptual and motor control literature (Todorov and Jordan, 2002; Wolpert et al., 1995) , but have been rarely applied to joint action (Wolpert et al., 2003). We specifically argue that individuals involved in a joint action builds an internal representation of their partner's ongoing actions and/or intentions (partner model) and use it to establish an interaction (Chackochan and Sanguineti, 2019) . The proposed AP architecture can be used to compare actual (human) and idealized (artificial) partners in order to test different assumptions on how we develop a coordination. We assessed the AP performance in an interaction paradigm where two agents are mechanically connected through a virtual spring, and are instructed to perform reaching movements from the same start and target points, but through different via-points.

3.2 Modeling framework

In this section we provide the background ideas, and the empirical basis for the proposed AP architecture. We specifically focus on (i) optimality in perception and control; (ii) the need to represent partners in joint action; and (iii) the way a coordination strategy can be developed through repeated practice.

3.2.1 Optimality in perception and action

Sensorimotor control is the end result of two inter-related processes, i.e. estimation of the state of the body and the external environment, and movement control. Motor commands - muscle activations - are generated on the basis of our movement goals and our belief in the current state (position, velocity) of our own body and the external environment. Both state estimation and movement control can be understood in terms of optimality principles. There is indeed ample evidence that our nervous system exhibits a close-to-optimal performance in both sensorimotor integration (Wolpert et al., 1995) and movement control Todorov and Jordan (2002). Hence the combination of optimal estimation and optimal control is the ideal framework for the development of biomimetic artificial partners.

Sensorimotor integration is believed to optimally combine prior beliefs and sensory feedback - e.g. visual, acoustic and/or haptic - to minimize the prediction uncertainty (Wolpert et al., 1995). Prior belief about body and environment dynamics is believed to rely on an internal representation of the causal relationship between the motor command and the body and environment state – the notion of 'forward model' of the body – and between the body state and its sensory consequences – a model of the sensory system, or 'sensory model'.

Optimal control posits that humans aim at maximizing their subjective utility, i.e. a trade-off between task-dependent movement cost – in reaching movements, it may simply be the endpoint error – and the perceived effort associated to movement. Hence actions result from a subjective evaluation of their associated costs and/or benefits. This subjective aspect can be summarized by a cost function that incorporates the task goals and requirements and accounts for effort (Shadmehr and Krakauer, 2008; Todorov, 2004). It is natural to extend this framework to joint action (Braun et al., 2009; Chackochan and Sanguineti, 2019; Li et al., 2019).

3.2.2 Partner representation

Sharing information is a key determinant of successful joint action Vesper et al. (2017). Information exchange implies verbal or non-verbal communication. If the participants are physically connected, haptic channel a powerful source of information about the partner (Chackochan and Sanguineti, 2019; Ganesh et al., 2014; Takagi et al., 2018; van der Wel et al., 2011). More reliable information generally leads to more efficient interaction (Chackochan and Sanguineti, 2019). However, stronger coupling provide more information but makes control more difficult, whereas weaker coupling facilitates coordination but provides less reliable information about partner actions (Takagi et al., 2018).

During interactions, individuals not only observe and react to partner actions but they also need to predict the actions of their partner and their expected consequences (Knoblich and Jordan, 2003; Vesper et al., 2017). To this purpose, joint action crucially requires that participants develop an internal representation of their partner or, using terminology from the sensorimotor control field (Franklin and Wolpert, 2011), a 'partner model'.

3.2.3 Game theory and fictitious play

In a joint action, the subjective utility of a participant typically depends on the state of all partners involved. Game theory provides the analytic and computational substrate for the underlying control processes. Therefore, it can be seen as an extension of optimal control to multi-agent scenarios. A game involves two or more individuals whose interests are neither completely opposed, nor completely coincident. In some joint action scenarios, agents agree on a shared strategy – for instance, through verbal communication – before the action takes place, thus behaving as a collective (Bacharach, 1999). In other scenarios there is no explicit prior agreement on a shared strategy. In these situations, coordination emerges gradually as each agent collects information about their opponent's actions, their outcomes and possibly

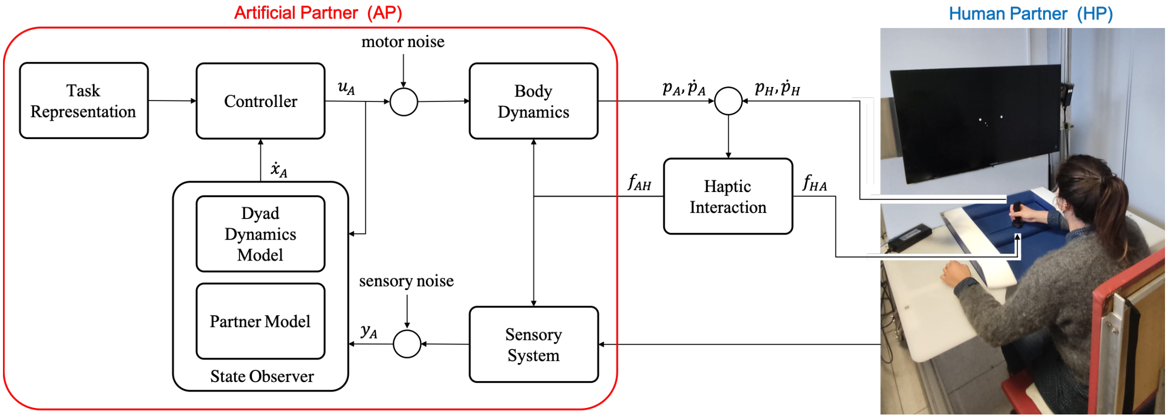

Figure 3.1 Artificial partner architecture. The artificial partner has a body and a sensory system. The control architecture includes a task representation, a feedback controller and a state observer, which also includes a partner model. This is the minimal architecture to support the development of an interaction. The control architecture is integrated into a robotic platform which supports the physical interaction with the human partner

their ultimate goals, using various mechanisms (Vesper et al., 2017) during continuous or repeated interaction. In motor control scenarios, differential non-cooperative games (Başar and Olsder, 1998) can model situations in which humans deal with their counterpart without speaking and by communicating just through sensory cues (visual, acoustic or haptic), but they independently determine their actions. It has been observed (Braun et al., 2009; Chackochan and Sanguineti, 2019; Jarrassé et al., 2012; Li et al., 2019) that when both agents have perfect information about the partner and the environment, they converge to a Nash equilibrium, i.e. a situation in which none of the players can unilaterally improve their benefit (Nash, 1951). When the players have competing goals, coordination develops gradually. They gradually gain knowledge about dyad dynamics, the task requirements, and partner's actions. Hence coordination (if any), is a result of learning and adaptation.

One possible solution of the problem of iteratively calculating a Nash equilibrium is represented by the classical learning process known as fictitious play (Berger, 2007; Brown, 1951). In fictitious play, two agents play the game repeatedly. After arbitrary initial moves in the first round, in every round each agent determines its best response against the empirical strategy distribution of their partner. Fictitious play posits that the players use a stationary strategy and only requires a model of the strategy distribution. However, the players do not have to know anything at all about their opponent's payoffs. All they do is to form beliefs about how their opponents will play (Fudenberg and Levine, 1998). Because of its minimal requirements, fictitious play has been proposed as a candidate mechanism for the development of optimal coordination in joint action, either involving movements (Chackochan and Sanguineti, 2019) or not (Grau-Moya et al., 2013).

3.3 Materials and Methods

3.3.1 Artificial Partner Architecture

The proposed artificial partner architecture directly derives from the computational model of joint action proposed in (Chackochan and Sanguineti, 2019). As in ’solo’ behavior, each motor command is generated by a feedback controller which relies on a state observer which predicts the overall dyad state. The only addition to the ’solo’ model is that, in order to correctly account for dyad dynamics, the state observer also needs to estimate the partner’s ongoing actions and/or movements. Based on the above computational model, a minimal but general AP architecture – see Figure 3.1 – should involve a task representation, a feedback controller, a state observer and a partner model. Interaction with its human counterpart also requires that the AP have its own body and sensory system. These can take different forms. The AP body may be physical – for instance, a humanoid robot – or virtual – an avatar moving in a VR or AR environment. Interaction may be visual – this requires that the AP is equipped with an artificial visual system – and/or physical – in this case, the AP may also involve a haptic interface, which exchanges interaction forces with the human player. Haptic channel provides additional information about partner movements, hence it is also part of AP’s sensory system. Although this AP concept is quite general, in this study we demonstrate it in a simplified situation. Both human and artificial partners move on a horizontal plane. The human partner (HP) grasps the handle of a planar robotic manipulandum, through which the AP and HP are physically coupled. Although very simple from the geometric and mechanical point of view, this scenario is central in many studies on computational motor control and in many rehabilitation robotics applications.

Dyad dynamics

We describe the dynamics of both the human and the artificial partner as two planar arms with two degrees of freedom (q_s for shoulder, q_e for elbow). We linearized arm dynamics around configuration $q_0 = [45^\circ, 90^\circ]$ and expressed it in Cartesian space:

$$\begin{aligned} M_H \ddot{p}_H + b_H \, \dot{p}_H &= f_H + f_{AH} \\ M_A \ddot{p}_A + b_A \, \dot{p}_A &= f_A + f_{HA} \end{aligned} \tag{3.1}$$

where p_A and p_H are the hand position vectors of, respectively, the artificial (A) and the human (H) partner and M_A and M_H are their respective inertia matrices. We set $M_i = J_i(q_0)^{-T} M_i(q_0) J_i(q_0)^{-1}$ with $i \in \{A, H\}$, where $J_i(q)$ is the jacobian of the forward kinematic transformation and $M_i(q)$ is the arm inertia matrix in joint space. Eq. 3.1 reflects the assumption that both partners are subjected to a small viscous force accounting for muscles and soft tissue damping. The interaction force f_{HA}:

$$f_{HA} = k \cdot (p_H - p_A) \tag{3.2}$$

is applied to the AP. An opposite force $f_{AH} = -f_{HA}$, is applied to the human partner through the robot handle. The position of the human partner, p_H, coincides with the position of the robot handle. As in (Todorov and Jordan, 2002), we modeled the dynamics of muscle force generation (f_A and f_H) as a second order system:

$$\begin{aligned} \tau_H^2 \ddot{f}_H + 2\tau_H \dot{f}_H + f_H &= u_H \\ \tau_A^2 \ddot{f}_A + 2\tau_A \dot{f}_A + f_A &= u_A \end{aligned} \tag{3.3}$$

where u_A and u_H are the activation vectors of human and artificial partners, which are taken as the dyad inputs, and τ_A, τ_H are the activation time constants (we set $\tau_A = \tau_H = 40$ ms). Eqs. 3.1, 3.2 and 3.3 can be expressed in state-space form by defining $x = \left[x_A^T \; x_H^T\right]^T$, as the state vector, where $x_i = [p_i^T \; \dot{p}_i^T \; f_i^T \; \dot{f}_i^T]^T$, with $i \in \{A, H\}$. The system has two control inputs, namely AP and HP muscle activations (u_A and u_H). In state-space form, Eqns. 3.1, 3.2 and 3.3 can be expressed as:

$$\dot{x} = \begin{bmatrix} A_{AA} & A_{AH} \\ A_{HA} & A_{HH} \end{bmatrix} \cdot x + \begin{bmatrix} B_A \\ 0 \end{bmatrix} \cdot (u_A + w_A) + \begin{bmatrix} 0 \\ B_H \end{bmatrix} \cdot (u_H + w_H) \tag{3.4}$$

where $w_A \sim N(0, \Sigma_A^w)$ and $w_H \sim N(0, \Sigma_H^w)$ are the 'motor'noise processes associated to, respectively, the artificial and the human partner, and A_i, A_{ij}, B_i, with $i, j \in \{A, H\}$, $i \neq j$, are defined as follows:

$$A_{ii} = \begin{bmatrix} 0_2 & I_2 & 0_2 & 0_2 \\ -k \cdot M_i^{-1} & -b_i \cdot M_i^{-1} & M_i^{-1} & 0_2 \\ 0_2 & 0_2 & 0_2 & I_2 \\ 0_2 & 0_2 & -I_2/\tau_i^2 & -2I_2/\tau_i \end{bmatrix} \tag{3.5}$$

$$A_{ij} = \begin{bmatrix} 0_1 & 0_2 & 0_2 & 0_2 \\ k \cdot M_i^{-1} & 0_2 & 0_2 & 0_2 \\ 0_1 & 0_2 & 0_2 & 0_2 \\ 0_1 & 0_2 & 0_2 & 0_2 \end{bmatrix} \quad B_i = \begin{bmatrix} 0_2 \\ 0_2 \\ 0_2 \\ I_2/\tau_i^2 \end{bmatrix} \tag{3.6}$$

where I_2 and 0_2 denote, respectively, the 2×2 identity and zero matrices.

Artificial partner dynamics and sensory system

Eq. 3.4 captures the dynamics of both the artificial and the human partner and their physical coupling. This general formulation will be used to simulate the artificial partner dynamics, which can be expressed as:

$$\dot{x}_A = A_{AA} \cdot x_A + A_{AH_1} \cdot p_H + B_A \cdot [u_A + w_A] \tag{3.7}$$

where A_{AH_1} denotes the first column of A_{AH}. For simulation and control purposes, the above equations were translated in discrete-time form:

$$x_A(t+1) = A^d_{AA} \cdot x_A(t) + A^d_{AH_1} \cdot p_H + B^d_A \cdot [u_A(t) + w_A(t)] \tag{3.8}$$

where A^d_{AA}, $A^d_{AH_1}$ and B^d_A are calculated from their continuous counterpart. To simplify the notation, in the following we will drop the 'd' but we will always refer to the discretized equations. The artificial partner's sensory system provides information about itself, its human counterpart (i.e., the whole dyad) and the external environment. Information about the human partner can be gathered both directly, e.g. through vision, thus observing p_H and $\dot{p}_H$, and through the interaction force f_{HA}. Hence the vector y_A of the sensory signals takes the form $y_A = [p_A, \dot{p}_A, p_H, \dot{p}_H, f_{HA}]^T$ and the sensory system is described as:

$$y_A = H_A \cdot x + v_A \tag{3.9}$$

The H_A matrix reflects the portion of sensory information which depends on system state:

$$H_A = \begin{bmatrix} I_2 & 0_2 & 0_2 & 0_2 & 0_2 & 0_2 \\ 0_2 & I_2 & 0_2 & 0_2 & 0_2 & 0_2 \\ 0_2 & 0_2 & 0_2 & I_2 & 0_2 & 0_2 \\ 0_2 & 0_2 & 0_2 & 0_2 & I_2 & 0_2 \\ -kI_2 & 0_2 & 0_2 & kI_2 & 0_2 & 0_2 \end{bmatrix} \tag{3.10}$$

The reliability of the sensory information is determined by the magnitude of the sensory noise, assumed to be Gaussian, $v_A(t) \sim N(0, \Sigma_A^v)$, with $\Sigma_A^v = \mathrm{diag}(\sigma_x^2 \cdot I_2, \sigma_{xd}^2 \cdot I_2, \sigma_x^2 \cdot I_2, \sigma_{xd}^2 \cdot I_2, \sigma_f^2 \cdot I_2)$. We set $\sigma_x = 1.7\,\mathrm{mm}$, $\sigma_{xd} = 35\,\mathrm{mm/s}$ and $\sigma_f = 2\,\mathrm{N}$.

State observer and partner model

Consistent with a large body of literature (Wolpert et al., 1995), we assume that the artificial partner optimally integrates sensory information its and own motor command (efferent copy) in order to estimate its own state, the state of its human partner, and possibly the state of the environment (for instance, target positions). Estimating human partner state is crucial in order to establish a successful interaction. While many studies agree that humans can form models of their opponents and/or they 'understand'their intentions, the exact nature of these models remains elusive. Interacting agents may simply estimate the ongoing partner movements. Or, they may additionally account for the partner's body dynamics, e.g. inertial properties, thus estimating the past motor commands (Chackochan and Sanguineti, 2019). Further, interacting agents may be able to infer their partner's control policy, i.e. the mapping between the state of both players and the partner's motor command (Li et al., 2019). This can be seen as a representation of the partner's ultimate goal or intentions (Sebanz et al., 2005). Whatever its form, estimating partner state requires a forward model of the human partner's dynamics and the availability of suitable sensory information – e.g. vision, proprioception, audition, etc. In the current AP implementation, the partner model fully accounts for the human partner's dynamic behavior. In particular, we assume that the state observer maintains a model of the whole dyad dynamics – which includes both artificial and human partner and their mechanical coupling. The state observer estimates the AP state x_A and the HP state, x_H. However, the latter is determined by the HP's motor command, u_H. Hence the state observer must be extended to predict the partner's motor command. Motor commands have sensory consequences, therefore the sensory consequences of movement carry information about the 'past' motor command. The complete set of 'prior'assumptions is summarized as:

$$\begin{aligned} x_A(t+1) &= A_{AA} x_A(t) + A_{AH} x_H(t) + B_A\left[u_A(t) + w_A(t)\right] \\ x_H(t+1) &= A_{HA} x_A(t) + A_{HH} x_H(t) + B_H\left[u_H(t) + w_H(t)\right] \\ u_H(t+1) &= u_H(t) + w_u(t) \end{aligned} \tag{3.11}$$

where $w_u(t) \sim N(0, \Sigma_x^u)$. The last equation reflects the prior belief that the human partner's input is an integrated Gaussian noise. In all experiments we set $\Sigma_x^u = 0.5\mathrm{N}^2$. The state observer is defined in terms of the augmented state vector $X_A = \left[x_A^T \; x_H^T \; u_H^T\right]^T$ as:

$$\begin{aligned}\hat{X}_A(t+1) = \; & A_A \cdot \hat{X}_A(t) + B_A \cdot u_A(t) + \\ & K_A(t) \cdot \left\{y_A(t) - [H_A \; 0]\hat{X}_A(t)\right\}\end{aligned} \tag{3.12}$$

Where A_A is defined as:

$$A_A = \begin{bmatrix} A_{AA} & A_{AH} & 0 \\ 0 & A_H & B_H \\ 0 & 0 & A_u \end{bmatrix} \tag{3.13}$$

The Kalman gain $K_A(t)$, $t = 1, \cdots, T$ of the innovation term reflects the trade-off of the reliability of the prediction and correction terms.

Task representation, optimal controller and fictive play

The AP task is specified by a quadratic cost functional:

$$\begin{aligned} J_A[u_A, u_H] = & x(T)^T \cdot Q_T \cdot x(T) + \\ & \frac{1}{T}\sum_{t=0}^{T-1} x(t)^T \cdot Q(t) \cdot x(t) + \\ & \frac{1}{T}\sum_{t=0}^{T-1} u_A(t)^T \cdot R(t) \cdot u_A(t) \end{aligned} \tag{3.14}$$

The above cost functional completely specifies AP's control policy if the human motor command is known.

Consistent with the fictitious play notion, we assumed that the HP's motor commands (dynamic partner model) or the movements (kinematic partner model) do not change from the previous trial. In other words, at each trial the AP 'sees'a plant that incorporates the partner's state estimated at the previous trial. This results in an affine linear dynamical system, and the resulting LQG controller has a feedback and a feedforward component:

$$u_A(t) = -L_A^{tr}(t) \cdot x(t) - l_A^{tr}(t) \tag{3.15}$$

In conclusion, the artificial partner autonomously determines its own control policy with no explicit agreement with the human counterpart, which in game theory is called a non-cooperative scenario. This implementation of fictitious play only uses the most recent estimate of partner's input. This is less robust than estimating the distribution of partner inputs over multiple repetitions, but may be adequate for practical purposes.

In this formulation of the controller movement duration is specified in advance. This constraint may be relaxed by including temporal discount terms in an infinite-horizon cost functional.

Temporal synchronization

Temporal synchronization is one basic feature of any joint action (Vesper et al., 2010) – first and foremost, partners synchronize the start of their movements. Entrainment - i.e. the tendency to fall on the same rhythm - has been observed in unintentional action or when people are not required to synchronize. However, in intentional joint action, entrainment cannot fully explain temporal synchronization as the latter requires mutual adaptation (Sebanz and Knoblich, 2009). To provide the artificial partner with the capability of synchronizing their movements with its human counterpart, we defined a simple mechanism which adapts the AP's start time, on a trial by trial basis, to the human start time:

$$TS^A(\mathrm{tr}+1) = TS^A(\mathrm{tr}) + \alpha \cdot \left[TS^H(\mathrm{tr}) - TS^A(\mathrm{tr})\right] \tag{3.16}$$

Where tr is the trial number, and TS^A and TS^H are the start times (measured from target onset) of, respectively, artificial and human partner. The artificial partner uses a simple threshold mechanism (speed grater than 0.1 m/s) to estimate $TS^A(\mathrm{tr})$ at each trial. Parameter α specifies the trade-off between stability and plasticity. In all experiments we set $\alpha = 1$, i.e. the AP sets its start time to that observed in the human partner on the previous trial.

Implementation

The AP architecture has been implemented with two distinct haptic interfaces: a planar serial manipulandum with two degrees of freedom (braccio di ferro, (Casadio et al., 2006)) and a planar cartesian manipulandum (H-MAN, Articares Ltd) (Campolo et al., 2014). In both cases, the implementation consists of two interconnected Simulink Desktop Real-Time and MATLAB applications. The Simulink application carries out real-time control of the artificial partner; the MATLAB application updates the optimal controller at the end of each trial on the basis of the estimated partner model. Within this implementation, the update rate of the artificial partner was 1 kHz. In the case of the H-MAN robot, the AP position was updated at 200 Hz. The refresh rate of the graphic display was 40 Hz and the data (position of both human and artificial partner and interaction forces) were saved at 100 Hz for further analysis.

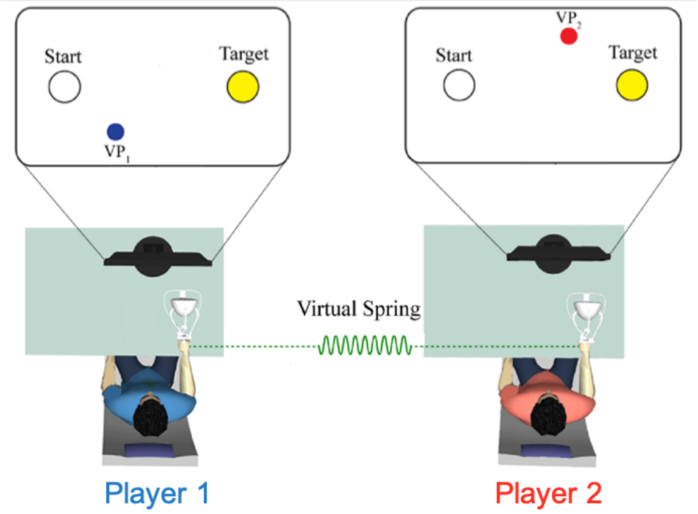

Figure 3.2 Two-via point joint action task. Each player has the same starting and end point but different via-points (VP_1, VP_2). Each player seat in front of a computer screen and has to grasp the handle of a 3D haptic interface (Novint Falcon). They are mechanically coupled by a virtual spring and instructed to perform planar point-to-point movements through different via-points. Each player can only see the final target and his/her own via-point. The players can not see each other or communicate verbally. Modified from (Chackochan and Sanguineti, 2019)

3.3.2 Experiments

Although very simple, the proposed AP architecture is suitable for use in all finite-horizon scenarios that involve planar movements in the horizontal plane. These include situations investigated in several experimental studies of joint action and many robot-assisted rehabilitation protocols. In this preliminary study, we focused on two aspects of the AP: (i) its ability to unveil the ongoing movements of the human partner, and (ii) its ability to adapt its behaviors in order to optimize the interaction. To test these functionalities, we focused on a joint action scenario in which two partners – mechanically connected through a virtual spring – are required to perform reaching movements from the same starting point to a common target, by crossing an intermediate via-point (VP) which is different for each subject. The players can only see their own VP but cannot see or speak to each other and they are instructed to keep the interaction force to a minimum, so that they must somehow coordinate in order to accomplish their respective goals. The task can be interpreted as a non-cooperative coordination game. A recent study (Chackochan and Sanguineti, 2019) reported that human players gradually develop a form of coordination which tends to the theoretical Nash equilibrium for that game – corresponding to both partners following the same trajectory through both VPs. The experimental setup and task are summarised in Figure 3.2.

In all experiments, the human participant sits in front of a computer screen. The subject grasps the robot handle and so that, when the hand is located approximately at the center of the workspace (origin of the reference system), the shoulder joint is flexed at about 45° with respect to the left-right shoulder direction and the forearm is flexed at about 90° with respect to the elbow. As in the human-human interaction experiments (Chackochan and Sanguineti, 2019), the subject can see the start and target point, its own via-point and their own hand position (a cursor on the screen), and is instructed to grasp the robot's handle with the right hand and to control the cursor motion, moving it from the start point to the target by crossing the via-point and keeping the interaction force low. Participants were not informed that he/she would have performed a joint task with a partner – either human or artificial. The AP has identical requirements, which are completely specified by the following cost function:

$$\begin{aligned} J_A[u_A, u_H] = & w_p \cdot \|p_T - p_A(T)\|^2 + \\ & w_v \cdot \|\dot{p}(T)\|^2 + \\ & w_{vp} \cdot \|p_{VP_A} - p(TC_A)\|^2 + \\ & w_f \cdot \frac{1}{T} \sum_{t=1}^{T} \|p_H(t) - p_A(t)\|^2 + \\ & r \cdot w_u \cdot \frac{1}{T} \sum_{t=1}^{T} u_A(t)^2 \end{aligned} \tag{3.17}$$

The cost function has five terms. The first two terms enforce stopping on target at the end of the movement. The third term reflects the requirement to pass through the via-point. The fourth term accounts for minimising the distance between agents throughout the movement. The last term penalises the effort incurred during the movement, i.e. it encourages the AP to keep the interaction force low. The weight coefficients determine the relative importance of the corresponding constraint. We set these weights by assuming (Bryson's rule) a maximum acceptable displacement (in the via-point and in the final target) equal to, respectively, the radius of the via-point ($r_T = 2.5$ mm) and that of the target ($r_{VP} = 5$ mm). Similarly, we calculated the value of the 'velocity' weight by assuming a maximum acceptable speed at the target of $d_v = 5$ mm/s. We made a similar normalisation in the maximum inter-agent distance ($r_{AH} = 15$ mm) and maximum activation ($u_{max} = 10$ N). In all experiments we used the following weights: $w_p = 1/r_T^2$, $w_{vp} = 1/r_{VP}^2$, $w_v = 1/d_v^2$, $w_f = 1/r_{AH}$ and $w_u = 1/u_{max}^2$. The scalar coefficient r - which specifies the trade-off between task-related accuracy and effort - was set to 1. Greater r implies a greater sensitivity to effort and therefore lower controller gain. The time of crossing of the via-point is also part of the optimization. Based on some prior evidence that VP crossing times are approximately proportional to the fraction of path length covered at the time of crossing (PL), at each trial we set the crossing time as $TC^{tr+1} = PL^{tr} \cdot MT_A^{tr}$, where MT_A^{tr} is the total movement duration.

Experiment 1 – Partner model

We expect that AP's ability to correctly estimate HP movements is affected by the strength of the physical connection and by the reliability of the haptic channel. We systematically varied the stiffness of the virtual spring k and the variance σ_f^2 of the haptic sensory channel. During this experiment, the human subject could see both via-points and was instructed to perform reaching movements from the start to the end point, by crossing both VPs. The robot force was switched off, i.e. $f_{AH} = 0$. In this way, the human subject was free to move with no AP intervention, but physical connection with the partner was always present on the AP side (so that $f_{HA} \neq 0$). The AP was programmed to make reaching movements through their own VP according to the cost functional of Eq.3.17. The simulated AP sensory system provided information about own movements, own VP, the target and the interaction force with the human subject, f_{HA}, but no direct information about position and velocity of the human subject. In summary, the human subject performs unperturbed reaching movements through two VPs but does not participate in the interaction. In contrast, the AP aims at adapting its movements in order to keep its goals (crossing its own VP) and to minimize the interaction force. As HP performance is almost stationary, accomplishing the task by the AP solely depends on its ability to predict own and partner actions. The experiment was organized in 30 epochs of 10 trials each. Each epoch corresponded to a different combination of stiffness and noise variance, in random balanced order. Specifically, we used six stiffness values: $100, 150, 200, 300, 400, 500$ N/m and five noise standard deviation values: 0.1, 1, 2, 3, or 4 N ($6 \times 5 = 30$). Only one human subject (28 y, female) participated in this experiment as the focus here was on AP performance.

Experiment 2 – Learning coordination

The goal of this experiment was to test various aspects of AP performance, in particular its capability to develop coordinated movements with a HP and whether they resemble those observed in two human partners. The experimental protocol was identical to (Chackochan and Sanguineti, 2019). Mechanical connection between AP and HP was now bi-directional, so that their movements were inter-dependent. Both AP and HP could only see their own via-point, and the only information about their partner was provided by the interaction force. The experiment was organized in 13 epochs of 12 trials each (a total of 156 trials). In the first epoch (baseline phase), mechanical coupling was turned off, and AP and HP acted alone. During epochs 2-11 the partners were mechanically connected. During epochs 12-13 (aftereffect phase) the force was removed again. At the end of each trial, the human subject was provided a 0-100 score reflecting their performance (a combination of distance to own via-point and average interaction force - the lower the better). The experimental session lasted approximately one hour. The only difference with respect to the original experimental protocol was an additional control on movement time. This was necessary because of the finite-horizon implementation of the current AP ($MT = 2.5s$ in the current experiment). If the human participant's movement time was $2s < MT < 3s$, at the end of the trial the target color turned green (appropriate duration); otherwise it turned red (wrong duration). A total of four subjects (25-27 y, 3M+1F) participated in this experiment. Their performance was compared with the data from five human-human dyads (Chackochan and Sanguineti, 2019), haptic (H) condition (25 ± 5 y, 9 M + 1 F).

The research conforms to the ethical standards laid down in the 1964 Declaration of Helsinki and was approved by the competent ethical committee (Comitato Etico Regione Liguria). Each participant signed a consent form conforming to these guidelines.

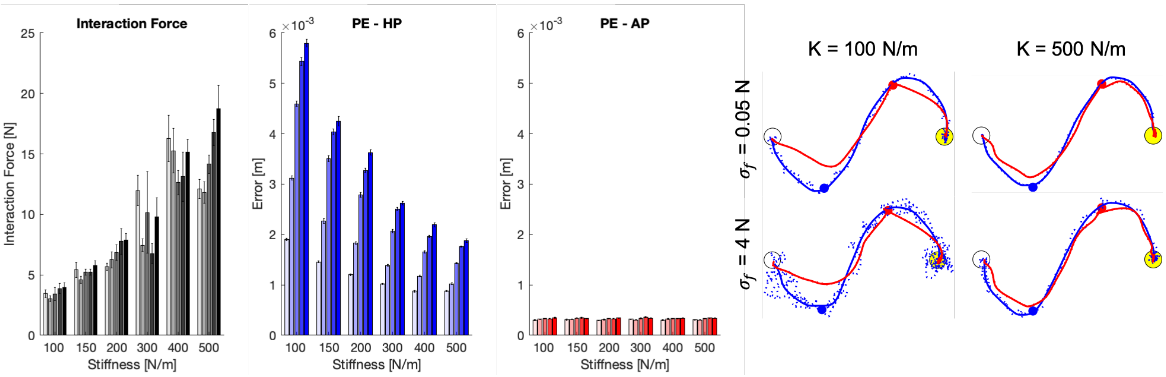

Figure 3.3 Experiment 1 - Partner model. Left: The interaction force increases with stiffness. Each bar corresponds to an individual epoch and therefore to a specific stiffness-noise pair. The bars were reordered for increasing stiffness and noise values. Noise magnitude is denoted by grey level (light grey: $\sigma_f = 0.5$ N; dark grey: $\sigma_f = 4$ N). Middle: Trajectory estimation accuracy (PE),for HP (red) and AP (blue). Data are reordered as for interaction force. All error bars denote standard errors. Right: AP (red) and HP (blue) trajectories in selected subject and trials, with different stiffnesses (K) and different noise standard deviation value σ_f. The continuous and dashed lines denote, respectively, the actual trajectories and the corresponding AP predictions.

3.3.3 Data Analysis

The recorded AP and HP movements and forces were smoothed and differentiated (4th order Savitzky-Golay, window size 370 ms).

Experiment 1

AP and human player are connected through a virtual spring. As a measure of interaction, for each trial (tr) we calculated the mean interaction force:

$$IF(tr) = \frac{1}{T(tr)} \sum_{t=0}^{T(tr)-1} |f_{HA}(t)| \tag{3.18}$$

where $T(tr)$ is the number of samples at trial tr. For each combination of stiffness and haptic noise magnitude, we took the IF average, standard deviation and coefficient of variation.

For each trial tr, we also computed the trajectory error PE between the actual and predicted trajectory of both the human and the artificial partner:

$$PE_i(tr) = \sqrt{\frac{1}{T(tr)} \sum_{t=0}^{T(tr)-1} |p_i(t) - \hat{p}_i(t)|^2} \tag{3.19}$$

where $p_i(t)$ is the position of player i, with $i \in \{A, H\}$, and $\hat{p}_i(t)$ is the position predicted by the AP's state observer. Likewise, we calculated the speed error SE between the actual and predicted speed profile of both the human and the artificial partner:

$$SE_i(tr) = \sqrt{\frac{1}{T(tr)} \sum_{t=0}^{T(tr)-1} |v_i(t) - \hat{v}_i(t)|^2} \tag{3.20}$$

where $v_i(t) = |\dot{p}_i(t)|$ is the speed profile. To examine the dependence of the above indicators on stiffness and noise level of the haptic channel, we used a repeated-measures ANOVA with two factors (stiffness and haptic noise magnitude). Due to the stochastic nature of AP behavior, we took each individual repetition of each condition ($N = 10$ trials per epoch) as an independent sample of AP behavior. We also corrected (Bonferroni-Holm) for multiple comparisons.

Experiment 2

We first assessed whether AP's temporal synchronization mechanism induced a mutual adaptation of the starting times of the two players, similar to that observed in human dyads. To do this, we compared the start times over trials of both artificial (TS^A) and human (TS^H) partners with those observed in previously reported human dyad experiments (Chackochan and Sanguineti, 2019). To analyze the overall dyad performance and convergence to joint coordination, we also evaluated the time course of spatial variability and of the minimum via-point distances.

Spatial variability was calculated as in (Liu and Todorov, 2007). We re-sampled all movements for a given subject at a fixed number of points equally spaced along the path, and found the average trajectory. Then, for each point along the average trajectory, we found the nearest sample point from each individual trajectory. These nearest points were averaged to recompute the corresponding point along the average trajectory, and the procedure was repeated until convergence. Hence this is a measure of path spatial variability, independent of time fluctuations.

The minimum via-point distances, MD_{ij}, were defined as the minimum distances at which player i gets closest to the j-th via-point.

To examine the dependence of the above indicators on training, we used paired-sample t-test comparing the indicators averaged over the initial and final epoch of the connected phase - epoch 2 and 11 respectively. As in Experiment 1, we corrected (Bonferroni-Holm) for multiple comparisons.

3.4 Results

3.4.1 Experiment 1 - Partner model

We first assessed the AP ability to predict its own movements and those of the human partner under various conditions (stiffness and noise level of the haptic channel). Figure 3.3 summarizes the partner model performance over conditions. As expected, the interaction force increases with stiffness; see Figure 3.3 (left). The observation was confirmed by statistical analysis. We observed a significant effect of stiffness ($p < 0.0001$), but no effect of noise or stiffness × noise interaction. The AP's state observer reliably predicts its own movements. As expected, prediction of the trajectory of the HP is significantly less reliable (greater PE: 0.25 ± 0.13 cm for HP vs 0.03 ± 0.0018 cm for AP); see Figure 3.3 (center). Stiffness and noise level affect prediction accuracy. As regards HP trajectories, the

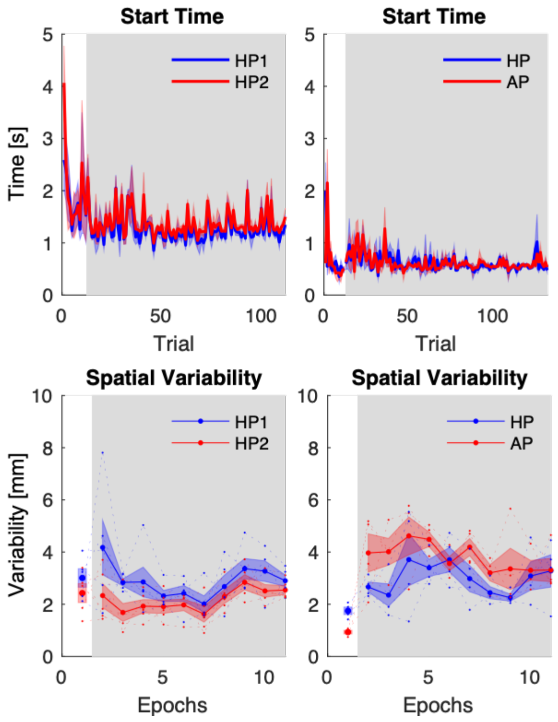

Figure 3.4 Top: Artificial and human partner gradually learn to synchronize, i.e. to start their movements at the same time (left) in a way that is qualitatively similar to H-H dyads (right). Bottom: temporal evolution of spatial variability over epochs. The artificial partner (right) is in red. H-H data taken from (Chackochan and Sanguineti, 2019). All panels report mean ± SE.

prediction error decreases significantly as stiffness increases ($p < 0.0001$) and noise decreases ($p < 0.0001$). We also found a significant stiffness × noise interaction ($p < 0.0001$). The AP prediction error increases as the sensory noise increases ($p < 0.0001$), but is not affected by stiffness. Similar results were observed for speed. The prediction accuracy is less reliable for HP (speed prediction error, SE, is about 5.76 ± 2.25 cm/s for HP vs 0.72 ± 0.22 cm/s for AP). The speed prediction error (SE) for HP decreases as the stiffness increases ($p < 0.0001$) and the noise level decreases ($p < 0.0001$). The error also increases with both stiffness ($p < 0.0001$) and noise level ($p < 0.0001$) for AP's speed estimation . We also observed a significant stiffness × noise interaction for both HP ($p < 0.0001$) and AP ($p = 0.0014$).

3.4.2 Experiment 2 - Learning coordination

The ability to synchronize the timing of the movements is crucial for the development of collaborative strategies Vesper et al. (2017). Figure 3.4 (top) compares the effect of AP-HP (A-H) synchronization with corresponding observations from human-human dyads.

Synchronization is qualitatively similar in A-H and H-H dyads. A-H and H-H also exhibit similar amounts of inter-trial variability; see Figure 3.4 (bottom).

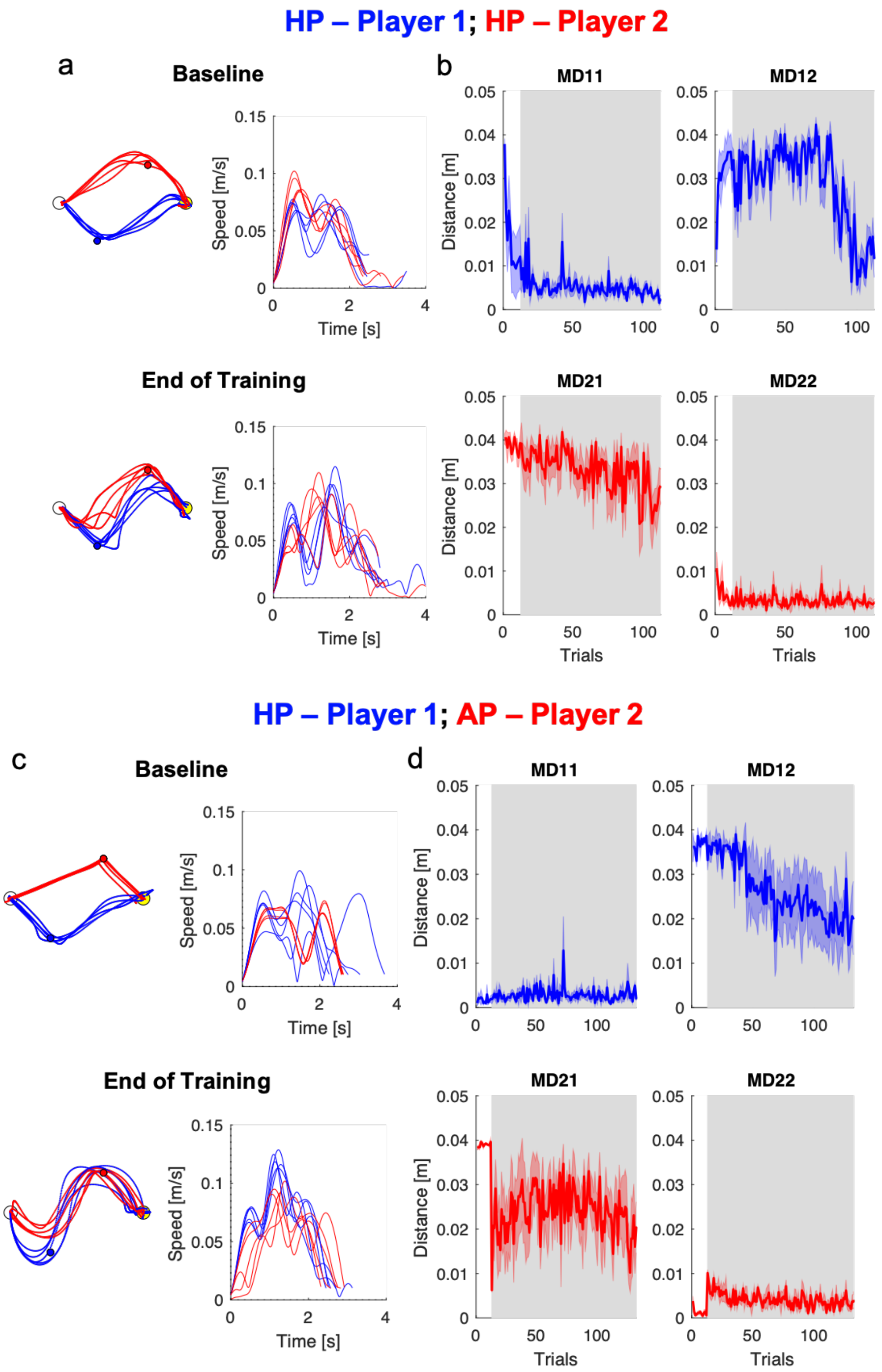

Figure 3.5 Trajectories and speed profiles during baseline and the last training epoch, for typical H-H (a) and A-H (c) dyads. In the A-H dyad, AP is in red, HP in blue. Both H-H and A-H dyads gradually develop a coordination. Temporal evolution of the minimum distance from via-point in human-human (b) and human-artificial (d) partner interaction (respectively, in blue and in red).

We then looked at the trajectories and speed profiles that the A-H dyad exhibits over repeated trials. During the baseline phase, AP and HP are not mechanically coupled and they simply perform reaching movements through their respective via-points; see Figure 3.5 a,c (top). The A-H and H-H dyads exhibit movements that are qualitatively similar in path, speed profile and inter-trial variability. During the adaptation phase, when mechanical coupling was turned on, both partners in a dyad gradually tend to move along similar paths by crossing their VP and getting closer to their partner's; see Figure 3.5 a,c (bottom). Again, the behavior of A-H dyads is qualitatively similar to their H-H counterparts.

In H-H dyads, both subjects decrease their distance from their opponent's via-point and maintain the distance from own via-point low along all epochs; see Figure 3.5 b. This indicates that both players prioritize their own goals and gradually improve coordination with their partners; see (Chackochan and Sanguineti, 2019). This behavior is a signature of the development of a coordination.

The same behavior is observed in A-H dyads. Over trials, both the AP and the HP gradually decrease their distance from their opponent's via-point, while at the same time keeping the distance from their own via-point low; see Figure 3.5 d. During interaction the HPs behave as in H-H dyads, by significantly decreasing their minimum distance to their partners ($p < 0.0001$); in APs the improvement is only partial (no significant decrease). At population level, we also found no significant change in start time and spatial variability.

3.5 Discussion

We presented a general computational framework for the development of artificial partners which are capable of establishing a coordination with a human partner. The architecture involves three main components: (i) a simulated body and sensory system, (ii) a state and partner observer, which includes an internal representation of the partner (partner model) and (iii) a feedback controller, based on a representation of the task in terms of a quadratic cost function. This formulation is rooted on probabilistic (bayesian) sensorimotor integration (Wolpert et al., 1995) and optimal control (Todorov and Jordan, 2002). We extend these principles – summarized in section II.A – in two directions. First, we argue that during joint action each participant infers their partner's intentions and/or ongoing actions see II.B (Curioni et al., 2019; Vesper et al., 2017). Building upon an experimentally confirmed model (Chackochan and Sanguineti, 2019), we formulate a sensorimotor integration model which predicts not only the dyad state, but also the ongoing partner actions. Second, consistent with the optimal control framework, we propose that coordination strategies emerge gradually through a simple adaptive process (fictitious play) which, in the case of perfect information, leads to a Nash equilibrium (Fudenberg and Levine, 1998). We present a specific implementation, which focuses on planar hand movements with finite duration, where the human and artificial partner are mechanically connected through a haptic interface. Under these assumptions, dyad dynamics and their sensory systems are described by linear time invariant dynamical models, with Gaussian noise. The artificial partner's task is described by a quadratic cost functional. Although simple, this formulation is applicable to a variety of experimental joint action scenarios, and its predictions can be directly compared to the observed human-human interaction outcomes.

3.5.1 The artificial partner estimates the actions of its opponent

We tested AP's ability to estimate the actions of a human opponent. As expected, we found that estimation improves as the strength of the coupling increases and as the noise variance decreases; see Figure 3.3. In the current implementation the AP keeps a dynamic model of the human, which fully accounts for its dynamic behavior. This is motivated by our previous report that players simply estimate the ongoing partner actions (Chackochan and Sanguineti, 2019). Other studies suggested that players develop more general partner representations, also accounting for intentions and ultimate goals Sebanz et al. (2005). Humans are indeed very good at extrapolating higher order information by observing the motion of their peers (Zunino et al., 2020). Alternative partner model formulations – for instance, the AP could use a 'kinematic' representation which only accounts for some specific features of human movements, like smoothness, but does not explicitly account for the underlying mechanisms. APs could serve as general tools to investigate what we represent about a partner and how we develop that representation.

3.5.2 Artificial and human partner learn to coordinate

Consistent with previous work (Chackochan and Sanguineti, 2019; Ganesh et al., 2014), we found that mechanical coupling can be exploited by interacting partners to exchange information and achieve a certain degree of coordination. Players in both A-H and H-H dyads tend to synchronise their start times and exhibit similar amounts of inter-trial variability; see Figure 3.4.

In H-H dyads, over trials the participants gather information about dyad dynamics and about the partner and adjust their actions accordingly. In the 2-via point task, this is reflected in the gradual decrease of the trajectory distances from both via-points; see Figure

??3.5. Although very simple, fictitious play reproduces this behavior in A-H dyads. Human participants interacting with the AP develop a coordination which is very similar to that observed in human-human dyads. However, APs converge less often to the trajectories corresponding to Nash equilibria. In fact, the observed trajectories suggest that the AP does not fully compensate for the HP. A systematic exploration of this issue – for instance, by systematically varying stiffness and sensory noise levels – is beyond the scope of this study which only aims at a technical validation of the basic AP architecture.

3.5.3 A general platform to study joint action

The proposed experiments focus on a specific sensorimotor interactive task, but the architecture and underlying assumptions – see Figure 3.1 – can be extended to investigate different tasks and scenarios. AP behavior is completely specified by a set of parameters, related to body dynamics (A_{AA}, A_{AH} and B_A), sensory system (H_A), perceptual and motor uncertainty (Σ^v_A, Σ^w_A), assumptions on partner model (dynamic, kinematic, or other), and personal traits like vigor (r) and willingness to establish a coordination. Other personal traits like risk sensitivity (Grau-Moya et al., 2012) can be easily incorporated. selection is crucial for determining the AP attitude toward the interaction. For instance, manipulating the weights of the 'effort' component of the AP cost function would lead the AP to prefer more or less effortful movements (and therefore smoother trajectories). Further, manipulating the start time adaptation rate (parameter α) would control AP's ability to synchronize with the human partner. The use of realistic APs with different attitudes and personal traits toward the interaction simplifies the study of joint action. In the A-H experimental paradigm one mind is a completely specified, thus allowing to characterize how the human counterpart responds.

3.5.4 Artificial partners as diagnostic and rehabilitation tools

Biomimetic partners with a realistic, predictable behavior may be used to characterize the ability of a human partner to establish an interaction (Dumas et al., 2014b).The inherently human capability to build representations about the partner and integrate them with own internal representation is altered in some pathologies, e.g. autism spectrum disorders or schizophrenia (Izawa et al., 2016), and a better understanding of these alterations would help improving the diagnosis and suggest ways to contrast their consequences. However, APs may be useful not only to investigate joint action. They can also be used to facilitate skill learning and neuromotor recovery. Rehabilitation robots have been often described as 'artificial therapists' and many of the developed technological solutions are somehow inspired by the observed mechanics of therapist-patient interaction. Patient and therapist constantly exchange information during a session. Looking at their actions and at the response to micro-perturbations, the therapist gradually develops an understanding of patient impairment and recovery potential. In the case of rehabilitation, interaction should aim at maximising recovery, making it faster and more durable. When interacting with assistive robots, humans tend to incorporate assistive forces in their motor plan, thus reducing their active contribution to movement, which may have adverse effects on recovery (Emken et al., 2007). To counteract this, several heuristic mechanisms have been proposed to provide 'assistance-as-needed' (Marchal-Crespo and Reinkensmeyer, 2009). A robot with an inherent ability to develop optimal forms of interaction would provide assistance 'as needed' by design. Also, it would automatically adapt to patient recovery.

3.6 Conclusions

The proposed AP architecture is intended as a general modeling framework. The current formulation makes a number of specific assumptions – on representation of the partner's actions, on the way coordination is achieved – which clearly require further empirical examination. To this purpose, the model may constitute a valuable experimental tool to test different hypotheses. The presented model formulation is limited to planar arm movements and linear dynamics, but can be extended to more complex, higher dimensional non-linear scenarios, for which mathematical tools and numerically tractable implementations are now available (Todorov et al., 2012).

Chapter 4

Computational joint action: dynamical models to understand the development of joint coordination

4.1 Introduction

Many of our everyday activities take place in social settings and are coordinated with other people. Even seemingly simple interactions, like a pair of workers sawing timber with back and forth movements, a couple executing complex steps on a dance floor, two children playing badminton, or providing physical therapy to a patient, require that two individual minds are somehow connected and their bodies coordinated (Sebanz et al., 2006). The characteristic feature of these interactions is that subjects influence each others' behavior through coupled sensorimotor exchanges within continuous action spaces, continuously in time, and possibly over repeated trials. Many of those interactions require active coordination, which is manifested by physical and cognitive responses related to explicit knowledge of the interacting partner. Peculiar types of joint action are those in which the two partners are physically coupled through a tool (Reed et al., 2006) or a pair of haptic interfaces which can be programmed to implement a variety of interaction modalities (Ganesh et al., 2014).

An early version of this chapter has been accepted as: De Vicariis, L., Bandini, G., Chackochan, V. T., Sanguineti, V. (2023). Computational joint action: dynamical models to understand the development of joint coordination. VIII Congress of the National Group of Bioengineering (GNB).

This latter scenario is particularly useful for quantifying various aspects of joint action performance because it allows to experimentally manipulate all aspects of the interaction.

During the last two decades, discrete-time dynamical models have been used to investigate sensorimotor learning and adaptation. These models have helped to clarify the underlying mechanisms: the role of the sensory prediction error (Scheidt et al., 2001; Thoroughman and Shadmehr, 2000); the interplay of implicit and explicit adaptation mechanisms (Hwang et al., 2006); and the evidence of multiple concurrent adaptation processes across different time scales (Smith et al., 2006). The same modeling framework has been used to assess how adaptation is altered in a number of neurological conditions, such as cerebellar degeneration (Smith and Shadmehr, 2005) and multiple sclerosis (Casadio et al., 2008), and to model neuromotor recovery through robot-assisted exercise (Casadio and Sanguineti, 2012). Cheng and Sabes (2006) have addressed the potential and the limitations of this class of models.

It is tempting to use the same approach to get more insight on how we develop coordination strategies while we interact with our peers. Few studies have addressed the perceptual and control mechanisms of joint action (Braun et al., 2009; Chackochan and Sanguineti, 2019; De Vicariis et al., 2022; Takagi et al., 2017). Dynamical models rely on optimality principles to account for perception and action selection within a single movement.

Action selection in joint coordination can be interpreted in terms of game theory. Many studies have reported that the sensorimotor behavior of physically coupled subjects converges toward a Nash equilibrium – a situation in which none of the players can unilaterally improve their payoff (Nash et al., 1950). However, the vast majority of previous studies addressing joint action within a game theoretic framework only focus on equilibrium situations (Braun et al., 2009, 2011). Very few studies (Chackochan and Sanguineti, 2019; Grau-Moya et al., 2013; Lindig-León et al., 2021b) have addressed the way joint coordination is negotiated and learned. One simple learning strategy when the players play repeatedly the game is that at every round each player determines their best response on the basis of their beliefs about how their opponents will play (fictitious play, FP) – see (Fudenberg and Levine, 1998). FP has been proposed as a candidate mechanism for the development of joint decision-making (Grau-Moya et al., 2013) or joint action in humans (Chackochan and Sanguineti, 2019), at least in situations where the players use stationary strategies. Reinforcement learning mechanisms been also implicated in situations where the players have an inaccurate knowledge about their task (Lindig-León et al., 2021b). The observation that players make errors during action selection has led to Quantal Response Equilibria (QRE) (McKelvey and Palfrey, 1995), which replace the perfectly rational expectations equilibrium embodied in Nash equilibrium with an imperfect, or noisy, rational expectations equilibrium. In this perspective, best response functions become probabilistic rather than deterministic. Uncertainty in action selection may also reflect partial information about own goal.

Perception is also believed to optimally combine prior beliefs and the available observations to minimize the prediction uncertainty (Körding and Wolpert, 2004; Wolpert et al., 1995). These same mechanisms likely mediate the prediction of the actions of our partners (Chackochan and Sanguineti, 2019). Very much like uncertainty in action selection, partial information about the partner affects the development of a joint coordination.

Here we show a general modeling framework for the development of coordination strategies in joint action. The model can be used to address a variety of interactive scenarios, both as a simulation and an analytical tool. We demonstrate the method in the context of sensorimotor versions of classical games with single (Prisoner's Dilemma) or multiple Nash equilibria (Stag Hunt). We show that the model reproduces the experimental findings (Prisoner's dilemma). We also speculate about how different partner representations may lead to different Nash equilibria (Stag Hunt). We then focus on a joint coordination game (Chackochan and Sanguineti, 2019), in which participants are mechanically coupled, have a shared final target, and are instructed to cross a via-point (VP, different for each player) while minimizing their interaction force. We estimate the model parameters from the experimental data (action time series over trials) and show that model correctly captures the different experimental conditions and the determinants of convergence to joint coordination.

4.2 General Computational Framework

We model a joint coordination task as a static non-cooperative infinite game (Başar and Olsder, 1998). We assume that players have incomplete information. In particular, each player (i) knows their own goal, but not their partner's, and (ii) has limited information about their partner's action. We also assume that learning of a coordination strategy occurs through fictitious play (FP) (Fudenberg and Levine, 1998), in which the agents play the game repeatedly, and at every round, each agent determines its best response on the basis of their beliefs about how their opponents will play. The overall model is depicted in Figure 4.2 using the graphical model formalism. At each trial t, the i-th player selects an action, $u_i(t)$s, on the basis of the partner's expected move (partner model), $x_i(t) = \hat{u}_i(t)$. Once the actions have been carried out, through their sensory systems both players collect information $y_i(t)$ about their own and their partner's actions. This information is then used to update their partner model, which will be accounted for when selecting the action on the next trial.

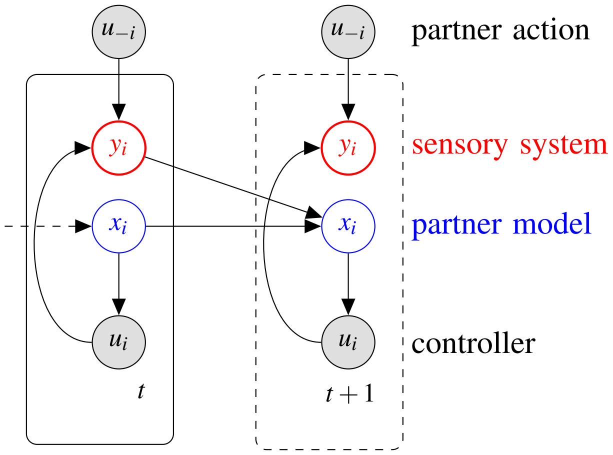

Figure 4.1 Computational model. Nodes in gray denote observed quantities, the other nodes denote quantities that are non-observed (latent). At each time step, each player's partner model (x_i) is used to select the actual action (u_i). Player i receives information about own and partner action (u_{-i}) through their sensory system (y_i). This information is used to determine the next action (u_i)

4.2.1 Joint tasks as quadratic games

For each player, i the task is specified by a quadratic cost function, which depends on both players' actions:

$$\begin{aligned} J_i(u_1,u_2) = \; & \tfrac{1}{2}\left[u_1^T R_{11}^i u_1 + 2u_1^T R_{12}^i u_2 + u_2^T R_{22}^i u_2\right] \\ & + r_1^{i\,T} \cdot u_1 + r_2^{i\,T} u_2 + z^i \end{aligned} \tag{4.1}$$

where $i = 1,2$ are the two players. Static games with infinite action space and quadratic costs are known as quadratic games. If all R_{jk}^i are symmetric and positive definite, the Nash equilibrium is uniquely determined (Başar and Olsder, 1998) as:

$$u^* = -R^{-1} \cdot r \tag{4.2}$$

where $u^* = [u_1^T \, u_2^T]^T$, $r = [r_1^1 \, r_2^2]^T$ and:

$$R = \begin{bmatrix} R_{11}^1 & R_{12}^1 \\ R_{21}^2 & R_{22}^2 \end{bmatrix} \tag{4.3}$$

Some joint action scenarios have multiple Nash equilibria, which implies that players have multiple strategy choices. To extend the quadratic game scenario to these situations, we assume that the cost incurred by player i is affected by the decision variable, $s_i \in \left\{S_i^1 \cdots S_i^M\right\}$,

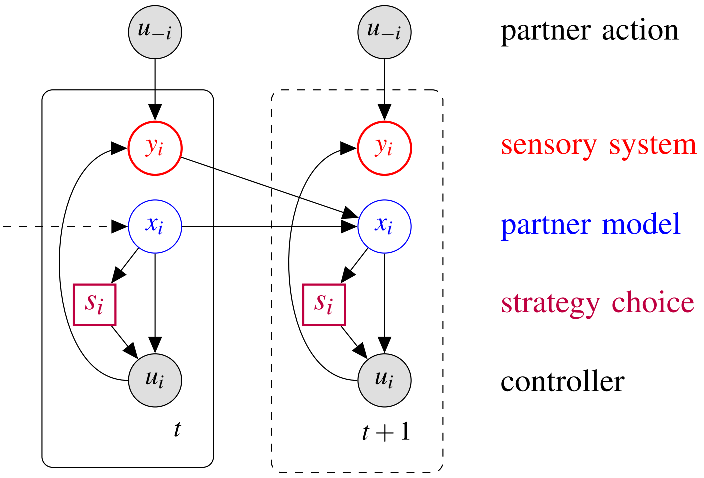

Figure 4.2 Computational model, extended to account for multiple strategy choices. At each time step, Player i receives information about partner action (u_{-i}) through their sensory system (y_i). This information is used to predict the next partner action (partner model, x_i). The partner model is used to select the actual action (u_i) which also depends on the strategy choice (s_i). The latter decision is also affected by the player's knowledge about the partner. For each player, sensory input (y_i), partner model (x_i) and strategy choice (s_i) are latent variables; actions (u_i and u_{-i}) are observable variables (denoted by shaded nodes)

which specifies one of the M strategies to be selected by the i-th player. Therefore Eq. 4.1 becomes:

$$\begin{aligned} J_i(u_1,u_2,s_i) = \; & \tfrac{1}{2}\left[u_1^T \cdot R_{11}^i(s_i)\cdot u_1 + 2\,u_1^T \cdot R_{12}^i(s_i)\cdot u_2 + u_2^T \cdot R_{22}^i(s_i)\cdot u_2\right] + \\ & r_1^i(s_i)^T \cdot u_1 + r_2^i(s_i)^T \cdot u_2 + z^i(s_i) \end{aligned} \tag{4.4}$$

The extended model is depicted in Figure 4.2. The partner model affects both action selection and strategy choice.

4.2.2 Sensory system

At every trial t, the sensory system of each player provides an estimate of own and partner's action on that trial:

$$y_i(t) = H^i \cdot u_{-i} + L^i \cdot u_i + v_i(t) \tag{4.5}$$

where $v_i(t) \sim N(0,\Sigma_y^i)$ is a zero-mean sensory (measurement) noise of the i-th player, and Σ_y^i is its covariance. Parameters H^i and L^i reflect the structure of the sensory system. In particular, H^i and L^i respectively relate partner and own action and their observed features to their sensory consequences.

4.2.3 Partner model

At every trial, t, each player optimally estimates their partner's next action by combining predictions and actual observations into a 'partner model'. Player i's prior belief of their partner action u_{-i} can be summarized as:

$$u_{-i}(t+1) = A^i \cdot u_{-i}(t) + w_i(t) \tag{4.6}$$

where $w_i(t) \sim N(0, \Sigma_x^i)$ is a process noise and A^i is a retention (memory) parameter. Parameters A^i and Σ_x^i reflect the player's prior belief on partner behavior – respectively, the extent to which partner behavior is stationary or erratic across trials. In particular, $A^i = 0$ denotes a completely erratic (memoryless) partner behavior. Conversely, $\Sigma_x^i = 0$ denotes a purely deterministic partner behavior.

We will assume that each player optimally combines prior estimates and sensory information about partner actions. Let $\hat{u}_{-i}^-$ and $\hat{u}_{-i}^+$ be the estimates of partner action, respectively before and after the sensory information is available, i.e. after the action has taken place (hereafter, 'prior' and 'posterior' estimates). The posterior estimate combines the prior estimate, sensory information and the efferent copy u_i^* of player action – see Eq. 4.15. The prior estimate propagates to the next trial by accounting for the player's retention rate. In conclusion, we have:

$$\begin{aligned} \hat{u}_{-i}^+(t) &= \hat{u}_{-i}^-(t) + K_i(t)\left[y_i(t) - H^i \hat{u}_{-i}^-(t) - L^i u_i^*(t)\right] \\ \hat{u}_{-i}^-(t+1) &= A^i \cdot \hat{u}_{-i}^+(t) + w_i(t) \end{aligned} \tag{4.7}$$

where $K_i(t)$ is calculated iteratively through the Kalman algorithm, which also provides time-varying estimates of the prior and posterior covariance matrices of $\hat{u}_{-i}^-$ and $\hat{u}_{-i}^+$, i.e. $P_i^-(t)$ and $P_i^+(t)$:

$$\begin{aligned} K_i(t) &= P_i^-(t) H^{i^T}\left[H^i P_i^-(t) H^{i^T} + \Sigma_y^i\right]^{-1} \\ P_i^+(t) &= \left[I - K_i(t) H^i\right] P_i^-(t) \\ P_i^-(t+1) &= A^i P_i^+(t) A^{i^T} + \Sigma_x^i \end{aligned} \tag{4.8}$$

Hereafter, we will refer to the prior estimate as the partner model: $x_i(t) \equiv \hat{u}_{-i}^-(t)$. Equations 4.7 may be merged into:

$$x_i(t+1) = A^i\left[I - K_i(t)\cdot H^i\right]\cdot x_i(t) + A^i \cdot K_i(t)\cdot y_i(t) - A^i \cdot K_i(t)\cdot L^i \cdot u_i^*(t) \tag{4.9}$$

4.2.4 Action selection and strategy choice

The optimal action of players i at time t, $u_i(t)$, is the one which minimizes the expected cost $\bar{J}_i(u_i) = E\{J_i(u_i, u_{-i})\}$, where the average is calculated over the partner's action, u_{-i}:

$$\bar{J}_i(u_i) = E\{J_i(u_i, u_{-i})\} = \int J_i(u_i, u_{-i}) \cdot p(u_{-i}) \cdot du_{-i} \tag{4.10}$$

When multiple equilibria exist, each player need to select one among these strategies, s_i. Players determines their next action u_i on the basis of the task and on their belief on partner state, x_i. This is achieved by minimizing the expected cost $\bar{J}_i(u_i, s_i) = E\{J_i(u_i, u_{-i}, s_i)\}$, which explicitly depends on the selected strategy s_i. The expectation is calculated over the partner's action, u_{-i}.

$$\bar{J}_i(u_i, s_i) = E\{J_i(u_i, u_{-i}, s_i)\} = \int J_i(u_i, u_{-i}) \cdot p(u_{-i}) \cdot du_{-i} \tag{4.11}$$

The partner model provides expected value $x_i(t)$ and covariance matrix $P_i^-(t)$ of the partner's action so that the expected cost can be expressed in terms of these quantities:

$$\begin{aligned} \bar{J}_i(u_i, s_i) = \; & \tfrac{1}{2} u_i^T \cdot R_{ii}^i(s_i) \cdot u_i + \left[x_i^T \cdot R_{i-i}^i(s_i) + r_i^i(s_i)\right] \cdot u_i + \\ & r_{-i}^i(s_i) \cdot x_i + z^i(s_i) + \tfrac{1}{2} \mathrm{tr}\left[R_{-i-i}^i(s_i) \cdot P_i^-\right] + \tfrac{1}{2} x_i^T \cdot R_{-i-i}^i(s_i) \cdot x_i \end{aligned} \tag{4.12}$$

which can be rewritten as:

$$\bar{J}_i(u_i, s_i) = \frac{1}{2} \left[u_i - u_i^*(s_i)\right]^T \cdot R_{ii}^i(s_i) \cdot \left[u_i - u_i^*(s_i)\right] + q(s_i, x_i) \tag{4.13}$$

where

$$q(s_i, x_i) = -\frac{1}{2} u_i^*(s_i)^T \cdot R_{ii}^i(s_i) \cdot u_i^*(s_i) + r_{-i}^i(s_i) \cdot x_i + z^i(s_i) + \frac{1}{2} \mathrm{tr}\left[R_{-i-i}^i(s_i) \cdot P_i^-\right] + \frac{1}{2} x_i^T \cdot R_{-i-i}^i(s_i) \cdot x_i \tag{4.14}$$

Therefore, optimal action u_i and strategy s_i are determined by minimizing the expected cost with respect to both the action u_i and the strategy s_i:

$$(u_i^*, s_i^*) = \arg \min_{(u_i, s_i)} \bar{J}_i(u_i, s_i) \tag{4.15}$$

To account for variability in platers' actions – i.e. players not behaving in a deterministic way, or not having a perfect knowledge of their cost function – we assume that the selection of both strategy and action is in fact a stochastic process. The expected cost can be interpreted as an energy term so that the conditional probability of u_i and s_i can be expressed as a Boltzmann probability density function:

$$p(u_i, s_i|x_i) \propto \frac{1}{Z} e^{-\frac{1}{\lambda^i}\bar{J}_i(u_i,s_i)} \tag{4.16}$$

where λ^i reflects the 'temperature' of action selection, with greater λ^i leading to greater randomness. This expression captures the property that actions and strategies are not completely specified by their expected cost. If the matrices $R^i_{ii}(S^i_j)$ are positive definite, the normalization factor is calculated as:

$$Z = \sum_{j=1}^{M} \int e^{-\frac{1}{\lambda^i}\bar{J}_i(u_i,S^i_j)} du_i = (2\pi)^{d/2}\sqrt{\lambda^i} \sum_{j=1}^{M} \frac{e^{-\frac{q(S^i_j)}{\lambda^i}}}{\sqrt{|R^i_{ii}(S^i_j)|}} \tag{4.17}$$

By substituting into Eq. 4.16 and marginalizing over s_i we get:

$$p(u_i|x_i) = \sum_j p(u_i, s_i = S^i_j|x_i) = \sum_j p(u_i|s_i = S^i_j, x_i) \cdot \Pr(s_i = S^i_j|x_i) \tag{4.18}$$

Hence the action u_i has a Gaussian Mixture distribution, in which the strategy prior and the strategy-specific means and covariances all depend on x_i.

The strategy prior, i.e. the probability to select strategy S^i_k given the partner model x_i, is calculated as:

$$\Pr(s_i = S^i_k|x_i) = \frac{e^{-\frac{1}{2}\log|R^i_{ii}(S^i_k)| - \frac{q(S^i_k,x_i)}{\lambda^i}}}{\sum_{j=1}^{M} e^{-\frac{1}{2}\log|R^i_{ii}(S^i_j)| - \frac{q(S^i_j,x_i)}{\lambda^i}}} \tag{4.19}$$

The action probability given the partner model x_i and the strategy $s_i = S^i_k$ has a multivariate Gaussian distribution, with mean

$$u^*_i(S^i_k, x_i) = -R^i_{ii}(S^i_k)^{-1} \cdot \left[R^i_{i-i}(S^i_k) \cdot x_i + r^i_i(S^i_k)\right] = H_c(s_i) \cdot x_i + h_c(S^i_k) \tag{4.20}$$

and covariance $\Sigma^i_u(S^i_k) = \lambda^i \cdot R^i_{ii}(S^i_k)^{-1}$:

$$p(u_i|s_i = S^i_k, x_i) = (2\pi)^{-d/2} \frac{1}{\sqrt{\lambda^i}} |R^i_{ii}(S^i_k)|^{1/2} \exp\left\{-\frac{1}{2\lambda^i} \cdot \left[u_i - u^*_i(S^i_k)\right]^T R^i_{ii}(S^i_k) \left[u_i - u^*_i(S^i_k)\right]\right\} \tag{4.21}$$

Like in the quantal response equilibria (QRE) model (McKelvey and Palfrey, 1995), strategy choice depends on the task and is modulated by partner model uncertainty – the dependence of

$q(s_i, x_i)$ on P_i^-, see 4.14. Also, randomness in action selection reflects the task requirements and is modulated by the temperature parameter. For stable learning to occur, the 'temperature' parameter λ^i should decrease over trials. In the model, we specify initial temperature, λ_1^i, and assume that the temperature changes gradually over trials according to the following relation:

$$\lambda^i(t+1) = a^i \cdot \lambda^i(t) \tag{4.22}$$

with $\lambda^i(1) = \lambda_1^i$, and $a^i < 1$ is the temperature decay rate.

In the above derivation we have made the critical assumption that the $R_{ii}^i(S_j^i)$ are positive definite. In this case, the probability distribution of actions – Eq.4.18 – is a Gaussian or a Gaussian Mixture. However, in competitive games the gain of one player corresponds to the loss of the other; in this setting the cost functions may become negative, which implies that matrix $R_{ii}^i(S_k^i)$ may no longer be positive definite. In this case, Eq.4.18 is no longer a Gaussian or a mixture of Gaussians. However, if Z in Eq 4.17 remains finite – for instance, if the action space is limited – Eq.4.18 will still define a valid action probability density function. In conclusion, the model is very general and can be used to describe any static game with continuous action space.

The model completely specifies one player's behavior in terms of seven parameters - Table 4.1: initial partner model (μ^i) and its covariance (P_0^i); retention rate (A^i) and process noise (Σ_x^i), which quantify the predictability of the partner model; sensory covariance (Σ_y^i), i.e. the reliability of the sensory system; initial temperature (λ_1^i) (exploration rate) and temperature decay rate (a^i).

Parameter	**Description**
μ^i	initial partner model value: is the initial partner action estimate
P_0^i	initial partner model covariance: quantifies the reliability of the initial partner model
A^i	retention rate: reflects the extent to which the partner model decays over trials
Σ_x^i	process noise: quantifies the belief about predictability of partner actions
Σ_y^i	sensory noise: quantifies the reliability of the sensory system
λ_1^i	initial temperature (exploration rate)
a^i	temperature decay rate

Table 4.1 Summary of model parameters

4.2.5 Model identification

This stochastic game formulation can be used as a generative model, to simulate the evolution of a particular game. Or, it can be used as an analytical tool: given a time series of players' actions, it is possible to calculate the model likelihood with respect to that dataset. We need to derive a fitting procedure for the above parameters for both players. The only data available is the time series of both players' actions, i.e.

$$D = \{u_1(t), u_2(t) | t = 1, \ldots, T\} \tag{4.23}$$

All other model variables, i.e. the partner model x_i, the sensory input, y_i, and the strategy choice, s_i, are not directly observable, i.e. are latent; see Figure 4.2. For parameter identification purposes, player i model can be reformulated as

$$\begin{cases} x_i(t+1) = & F^i(t) \cdot x_i(t) + G^i(t) \cdot u_{-i}(t) + w_i(t) \\ u_i(t) = & H_c^i(s_i) \cdot x_i + h_c^i(s_i) + \eta_i(t) \end{cases} \tag{4.24}$$

where the true partner action $u_{-i}(t)$ is the model's external input, $x_i(t)$ is the internal state and $u_i(t)$ and $s_i(t)$ are the outputs, and the noise terms $\eta_i(t) \sim N(0, \Sigma_u^i)$ and $w_i(t) \sim N(0, \Sigma_w^i)$ can be interpreted as measurement and process noise. At each time step, the strategy $s_i(t)$ can be estimated from the partner model $x_i(t)$ and the observation $y_i(t)$ as the strategy choice which maximizes the posterior probability:

$$\Pr(s_i = S_k^i | x_i, y_i) = \frac{p(y_i | s_i = S_k^i) \cdot \Pr(s_i = S_k^i | x_i)}{\sum_j p(y_i | s_i = S_j^i) \cdot \Pr(s_i = S_j^i | x_i)} \tag{4.25}$$

where $p(y_i | s_i = S_j^i)$ and $\Pr(s_i = S_j^i | x_i)$ are calculated from Eq. 4.19-4.21 and:

$$\begin{aligned} F^i(t) &= A^i \left[I - K_i(t) \cdot H^i\right] \\ G^i(t) &= A^i K_i(t) H^i \\ H_c^i(t) &= -R_{ii}^i(s_i(t))^{-1} \cdot R_{i-i}^i(s_i(t)) \\ h_c^i(t) &= -R_{ii}^i(s_i(t))^{-1} \cdot r_i^i(s_i(t)) \\ \Sigma_w^i(t) &= A^i K_i(t) \left[\Sigma_y^i + L^i \Sigma_u^i(S_k^i) L^{i^T}\right] K_i(t)^T A^{i^T} \end{aligned} \tag{4.26}$$

All the above matrices are time-varying. Eq. 4.24 can be rewritten in predictor form:

$$\begin{cases} x_i(t+1) &= F^i(t) \cdot x_i(t) + G_i(t) \cdot u_{-i}(t) + W^i(t) \left[u_i(t) - H_c^i(t) \cdot x_i(t) - h_c^i(t)\right] \\ P_i(t+1) &= F(t) \cdot \left[I - W(t) \cdot H_c^i(t)\right] \cdot P_i(t) \cdot F(t)^T + \Sigma_w^i(t) \\ u_i(t) &= H_c^i(t) \cdot x_i + h_c^i(t) + v_i(t) \end{cases} \tag{4.27}$$

where:

$$W(t) = P_i(t) \cdot H_c^i(t) \cdot \left[H_c^i(t) \cdot P_i(t) \cdot H_c^i(t)^T + \Sigma_u^i(t)\right]^{-1} \tag{4.28}$$

The model parameters $b^i = \left\{\mu^i, P_0^i, A^i, \Sigma_x^i, \Sigma_y^i, \lambda_1^i, a^i\right\}$ can be identified using a non-linear optimization method.

For a given b^i, we use Eq. 4.27 to estimate the sequence $\{x_i(t), P_i(t) | t = 1, \ldots, T\}$. We then calculate the expected (minus-)log-likelihood, expressed in 'innovation' form:

$$\begin{aligned} \mathscr{G}(b^i) = \; & -\tfrac{1}{2}\sum_{t=1}^{T} \log\left|H_c^i(t)\cdot P_i(t)\cdot H_c^i(t)^T + \Sigma_u^i(t)\right| \\ & -\tfrac{1}{2}\sum_{t=1}^{T}\left[u_i(t) - \hat{u}_i^-(t)\right]^T \cdot \left(H_c^i(t)\cdot P_i(t)\cdot H_c^i(t)^T + \Sigma_u^i(t)\right)^{-1} \cdot \left[u_i(t) - \hat{u}_i^-(t)\right] \end{aligned} \tag{4.29}$$

where $\hat{u}_i^-(t) = H_c^i(t)\cdot\hat{x}_i^-(t) + h_c^i(t)$. $\mathscr{G}(b^i)$ is the function to be minimized (Gupta and Mehra, 1974) – see also (Shumway and Stoffer, 1982).

4.3 Applications to joint scenarios

The model can be used either to simulate interactive settings or as an analytic tool to describe experimental observations. In this section, we test the model in a series of interaction scenarios. In a motor version of the prisoner dilemma, previously proposed by Braun et al. (2009), we discuss the effect of the different model parameters on the development of a joint strategy. We then use a continuous version of the stag hunt game (Yoshida et al., 2008) to clarify how the model deals with multiple Nash equilibria. Finally, we simulate the two via-points task (Chackochan and Sanguineti, 2019). Using experimental data from this study, we relate joint coordination parameters with model parameters.

4.3.1 Prisoner's dilemma

Braun et al. (2009) proposed a motor version of the Prisoner dilemma classical game, in which each subject controls the position of a cursor $p_1 = \left[p_{1x} p_{1y}\right]^T$ and $p_2 = \left[p_{2x} p_{2y}\right]^T$ in a 2D workspace. The cursor is continuously controlled within a single trial. Each participant was instructed to place their cursor within a horizontal target bar (placed at height p_T, same for both players). At each trial, p_T was made to vary between 5 and 20 cm. Participants started from anywhere within a starting bar (horizontal bar placed at $p_T = 0$ cm) and were required to make a forward movement (y-direction) to touch the target bar. They were free to touch it anywhere along its 15 cm width. For successful trial completion, the target bar had to be reached by both players within 1500 ms. The final x-position was taken as their action choice, i.e. $u_i = p_{ix}$ with $i = 1, 2$. During the movement, both players experienced a 1-dimensional spring attached to the starting bar. The spring constants depended on the lateral positions p_{1x} and p_{2x} of both players, where $u_i = p_{ix}$ corresponds to a normalized horizontal

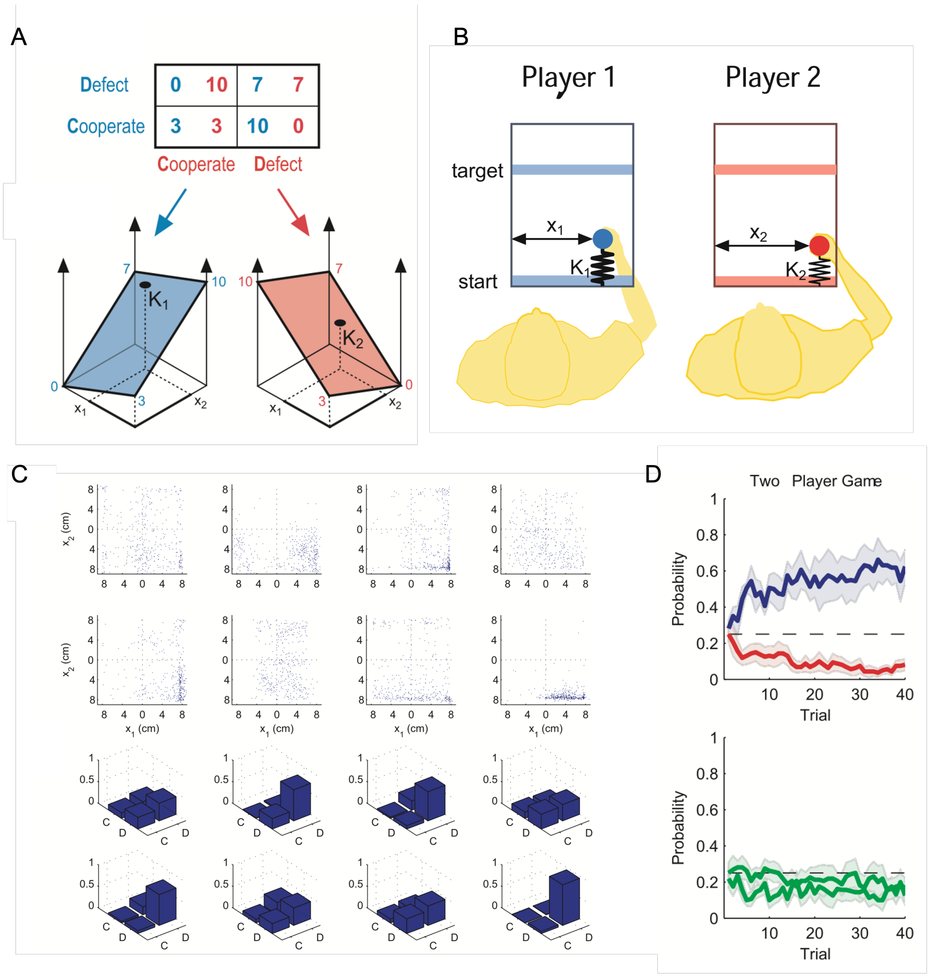

Figure 4.3 Prisoner dilemma motor game. The current figures are reprinted from Braun et al. (2009). A) Payoff matrix for the classical. prisoner dilemma game. B) Experimental setup of the motor version proposed. C) Experimental results are reported for each participating dyad in the dyadic action space and as the probabilities of the four different game solutions - i.e. Nash equilibrium, cooperative, and the two exploitative solutions.

position (ranging between 0 and 1). For each target bar, one edge was defined as defect (e.g. $p_{ix} = 0$) and the other as cooperate (e.g. $p_{ix} = 1$). In the experiment, the assignment of the defect/cooperation boundaries to the left/right side of each target was randomized. This gave four possible assignments (i.e. defect to left or right of target bar 1 and defect to left or right of target bar 2). Intermediate lateral deviations took on values between 0 and 1. The final x-position was categorized as 'cooperate' or 'defect' depending on whether $u_i > 0.5$ or $u_i < 0.5$. The spring constants were continuously updated as $k_1(u_1,u_2) = \alpha \cdot [3u_1 - 7(1-u_2)]$ and $k_2(u_1,u_2) = \alpha \cdot [3u_2 - 7(1-u_1)]$, with $\alpha = 0.19$ N/cm (constant over the whole experiment). These spring constants are linear interpolations of the classical prisoner dilemma matrix, with intermediate lateral deviations leading to intermediate spring constants. The participants experienced forces $F_1 = -k_1 p_{1y}$ and $F_2 = -k_2 p_{2y}$, resisting their forward motion. In the experiment, the participants performed 20 sets of 40 trials. At the start of each set the allocation of the target edges to defect/cooperate was randomized. Thus, within a set of 40 trials the same force landscape was applied.

The above task can be modeled as a quadratic game with the following cost function for each player i:

$$J_i(p_1, p_2) = \frac{1}{2} w_i |p_{iy} - p_T|^2 + \frac{1}{2} r_i |F_i|^2 = \frac{1}{2} w_i |p_{iy} - p_T|^2 + \frac{1}{2} r_i k_i (p_{1x}, p_{2x})^2 p_{iy}^2 \qquad (4.30)$$

Assuming that $p_{1y} = p_{2y} = p_T$, and taking $u_1 = p_{1x}$, $u_2 = p_{2x}$ the above simplifies as:

$$J_i(u_1, u_2) = \frac{1}{2} r_i \cdot p_T^2 \cdot \alpha^2 \cdot [4u_i^2 + 49u_{-i}^2 - 28u_i \cdot u_{-i} - 8u_i + 28u_{-i} + 4] \qquad (4.31)$$

In the simulations, we assumed that each player sees their own movement and the interaction force, i.e. $y_i = [u_i \; k_i(u_i, u_{-i}) p_T]^T + v_i$. In Braun et al. (2009), the players are not explicitly aware of their own task (cost function). In the model, this can be accounted for by specifying a greater initial temperature, λ_1^i. In Braun et al. (2009), the dyads performed ?? repetitions of the game. The analysis of the results focused on the temporal evolution of the probability to choose one among the four action pairs: cooperate-cooperate, defect-defect, cooperate-defect or defect-cooperate. We simulated the same experiment by systematically varying the sensory noise covariance Σ_y^i (low: 0.1; high: 2) and the temperature decay rate a^i (low: 0.9, high:1, i.e. complete retention). All other parameters remained constant. We set the partner predictability $\Sigma_x^i = 0.05$ and partner model retention $A^i = 0.99$. The initial temperature $\lambda_1^i = 1$ - was selected empirically in order to achieve – as in the actual experiments – approximately equal initial probabilities of selecting each of the four strategies.

4.3.2 Stag Hunt

In the classical Stag Hunt game, two hunters have the option of jointly hunting a stag, or independently pursuing a rabbit. Hunting a stag results in a greater payoff, but requires the cooperation of both partners. Catching a rabbit does not require cooperation, but gives a smaller payoff. This game has two pure-strategy Nash equilibria: one strategy (catching a rabbit) is risk-dominant with low-payoff states and can be attained without cooperation. The other strategy (catching a stag) is payoff dominant in the sense that both agents get a higher payoff, but only if they cooperate. The original version of the game assumed a discrete action

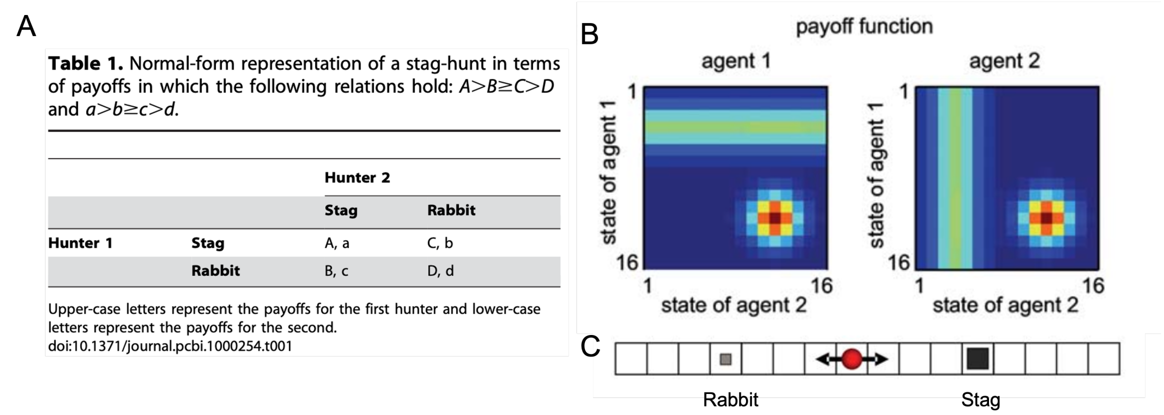

Table 1. Normal-form representation of a stag-hunt in terms of payoffs in which the following relations hold: *A*>*B*≥*C*>*D* and *a*>*b*≥*c*>*d*.

		Hunter 2	
		Stag	**Rabbit**
Hunter 1	**Stag**	A, a	C, b
	Rabbit	B, c	D, d

Upper-case letters represent the payoffs for the first hunter and lower-case letters represent the payoffs for the second.
doi:10.1371/journal.pcbi.1000254.t001

Figure 4.4 A continuous version of the Stag Hunt game. Figures in this panel are reprinted from Yoshida et al. (2008). A) Payoff matrix of the classical Stag-Hunt game. B) From the payoff matrix the payoff function is derived to span the range of actions of the two players. C) The motor task workspace. Players were initially equally distant from the two prey and were required to reach for one of the two (Stag/Rabbit).

space (rabbit, stag). Following Yoshida et al. (2008), we use a continuous state space with $u_i \in [1,16]$ for $i = 1,2$. Rabbit and stag do not move and are located at $u_R = 4$ and $u_S = 12$. The payoff for the stag is only achieved when both players are in u_S, whereas the payoff for the rabbit does not depend on the other agent's state. We introduce a discrete variable s_i which denotes whether player i is pursuing the 'rabbit' ($s_i = 1$) or the 'stag' strategy ($s_i = 0$). Therefore, the cost functions for the two players are:

$$\begin{aligned} J_i(u_i, u_{-i}, s_i) = \quad & s_i \cdot \tfrac{1}{2} w_R \cdot [|u_i - u_R|^2 + z_R] + \\ & (1 - s_i)\tfrac{1}{2} w_S [|u_i - u_S|^2 + |u_{-i} - u_S|^2 + z_S] \end{aligned} \tag{4.32}$$

The model parameters are identified so that the costs respect the cost matrix of the respective discrete game $a < b =< d < c$ and so that hunting for a stag is the payoff dominant solution and hunting a rabbit is the risk dominant solution. Yoshida et al. (2008) simulate scenarios in which the players do not account for their partner – hence the rabbit strategy is favored – and scenarios in which the partner is taken into account – in which case the stag strategy is favored. Therefore:

$$\begin{aligned}
&R^i_{ii} = s_i w_R + (1-s_i) w_S \\
&R^i_{i-i} = 0 \\
&R^i_{-i-i} = (1-s_i)\cdot w_S \\
&r^i_i = s_i w_R \cdot u_R + (1-s_i) w_S \cdot u_S \\
&r^i_{-i} = (1-s_i) w_S \cdot u_S \\
&z^i = \tfrac{1}{2} s_i w_R \cdot (u_R^2 + z_R) + \tfrac{1}{2}(1-s_i) w_S \cdot (2u_S^2 + z_S)
\end{aligned} \tag{4.33}$$

We simulated sequences of 1000 trials, in which we systematically varied the sensory noise covariance Σ^i_y (low: 1; high: 100) and the belief on partner predictability Σ^i_x (low: 1; high: 100). All other parameters remained constant. We set the partner model retention $A^i = 0.99$. The initial temperature λ^i_1 =1 and the temperature decay rate $\alpha^i = 0.99$ - were selected empirically in order to achieve approximately equal initial probabilities of selecting each of the four strategies.

4.3.3 2-VP task

In the two via-point (2-VP) task (Chackochan and Sanguineti, 2019), the players are mechanically connected through a virtual spring. They are instructed to perform reaching movement from the same start and end position, by crossing two different via-points. In addition, they are encouraged to keep the interaction force - and hence their distances – to a minimum.

The above assumptions can be summarized by the following cost function:

$$\begin{aligned}
J_i(p_1,p_2) = \;& \tfrac{1}{2} w_1 \left[p_i(t_i) - \mathrm{VP}_i\right]^T \cdot \left[p_i(t_i) - \mathrm{VP}_i\right] + \\
& \tfrac{1}{2}\tfrac{w_2}{T} \textstyle\sum_{t=0}^{T} \left[p_1(t) - p_2(t)\right]^T \cdot \left[p_1(t) - p_2(t)\right] + \\
& \tfrac{1}{2}\tfrac{r}{T} \textstyle\sum_{t=0}^{T} F_i(t)^T \cdot F_i(t)
\end{aligned} \tag{4.34}$$

where we take $t = 0$ as the start time and $t = T$ as the final time (common to both players) and we assume that $p_i(0) = p_{start}$ and $p_i(T) = p_{end}$. Time t_i is the time at which player i crosses their own via-point VP_i. We make the simplifying assumption that t_i is fixed across trials. Body dynamics of player i can be described as

$$m_i \ddot{p}_i(t) + k\left[p_i(t) - p_{-i}(t)\right] + b\dot{p}_i(t) = F_i(t) \tag{4.35}$$

where $F_i(t)$ is the motor command.

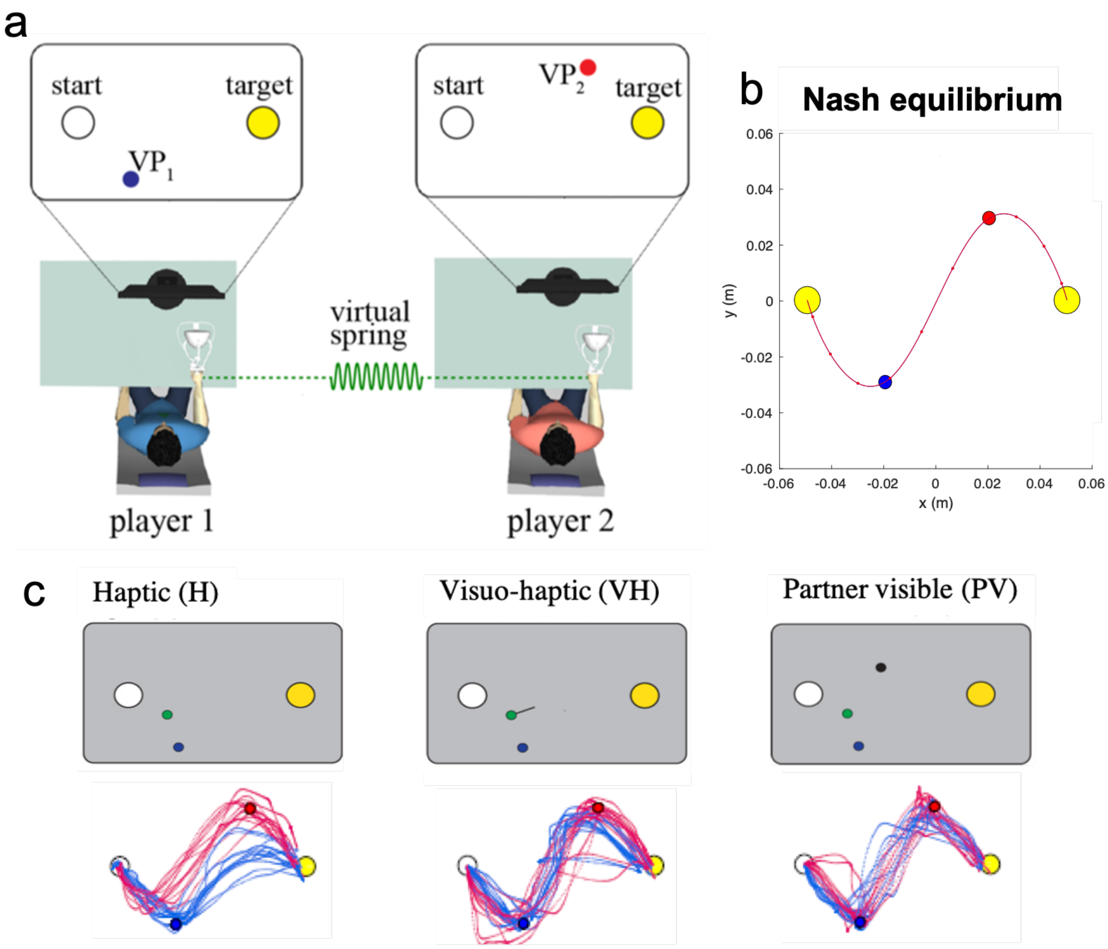

Figure 4.5 The 2-VP task. a: Partners in a dyad were connected through a virtual spring. They were instructed to perform reaching movements between the same start point and the same target point but through different via-points (VP). Each player could only see their own VP, but not their partner's. Both were instructed to keep the interaction force as low as possible during movement. b: Nash equilibrium for the 2-VP task. c: Trajectories at the end of training for one representative dyad for each group. Information about partner actions was provided haptically, through the interaction force (Haptic group, H), by additionally displaying the interaction force vector on the screen (Visuo-Haptic group, VH) or by also displaying the partner's cursor (Partner Visible, PV).

We approximate the whole trajectory $p_i(\cdot)$ for each player with a polyharmonic spline, specified by a small number M of nodes $u_i = [p_i(t_1) \; \ldots \; p_i(t_M)]^T$. The position at time t can be approximated as

$$\begin{aligned} p_i(t) = \; & \sum_{m=1}^{M} v_m^x |t-t_m|^3 + v_{M+1}^x \cdot t + v_{M+2}^x = \\ & g(t) \cdot v_i(t) = \\ & g(t) \cdot S \cdot u_i \end{aligned} \tag{4.36}$$

where $g(t) = \left[|t-t_1|^3 \; \ldots |t-t_M|^3 \; t \; 1 \right]^T$ and S is a constant matrix. Considering that some components of u_i^x (and u_i^y) are fixed, we finally have

$$p_i(t) = g(t)^T \cdot [S_1 \cdot u_i + S_0 \cdot u_0] \tag{4.37}$$

We define the 'action' of player i as the vector of nodes' coordinates, u_i. This approximation captures the overall shape of a generic trajectory in terms of a small number of nodes. Parameters w_1, w_2, and r are determined by the Bryson rule: $w_k = 1/d_{k_{\max}}$ where $d_{k_{\max}}$ is the maximum acceptable magnitude of the corresponding term of the cost function. In our task, we set $w_1 = 1/r_{VP}^2$, $w_2 = 1/d_{\max}^2$ and $r = 1/F_{max}^2$, where the VP radius is $r_{VP} = 2.5$ mm, the maximum between-subject distance is $d_{\max} = 1.5$ mm and the maximum force is $F_{max} = 0.1$ N. Using these parameters, the Nash equilibrium for this task is a joint trajectory in which both players cross both via-points - see Figure 4.5b. Further, in Eq. 4.5 we set $H^i = -L^i = I$ to reflect that what is sensed is the interaction force between the two players.

We estimated the model parameters from 120-trial time series of trajectories of 15 dyads (30 participants) – five dyads per group. Parameters describe the learning behavior of each player in all dyads. For each parameter we used a one-way ANOVA with Group as factor. In case of a significant group effect, we used post-hoc analysis to identify group differences.

4.4 Results

4.4.1 Prisoner's Dilemma

We used the model to simulate the experiment proposed in Braun et al. (2009). At each trial, the players' actions are defined by a scalar value, ranging between 0 (defect) and 1 (cooperate). Figure 4.6 summarizes the simulation results for each parameter combination. We report the overall distribution of actions and the time course of the probability to select one of the four possible strategies.

With all combinations of the parameters, the simulations exhibit a gradual increase of the probability to select the (defect, defect) strategy and conversely a gradual decay of the probabilities to select the other three. The probability rate of convergence is specifically affected by the temperature decay rate α^i – faster decay leads to faster convergence.

With lower sensory noise (Σ_y^i) the players tend to converge more to the Nash equilibrium. Instead, if players have higher values of sensory noise, actions tend to vary more in the action space. In particular, with reliable sensory information and a low exploration decay rate (α^i) players quickly converge toward the Nash equilibria. Consistently, with high exploration decay rate players still seems to act in the direction of the Nash equilibria, however, the convergence rate is much slower. In general, the model predicts that in order to converge to a Nash solution the players need to have reliable information about the partner and to be prone to explore possible more convenient solutions. This is consistent with the experimental results from Braun et al. (2009) where interacting agents tend to converge toward the Nash solution. Interestingly, the model is capable to generate data that resemble the experimental results but also makes predictions about possible other conditions that can be tested using the experimental setup proposed by Braun et al. (2009).

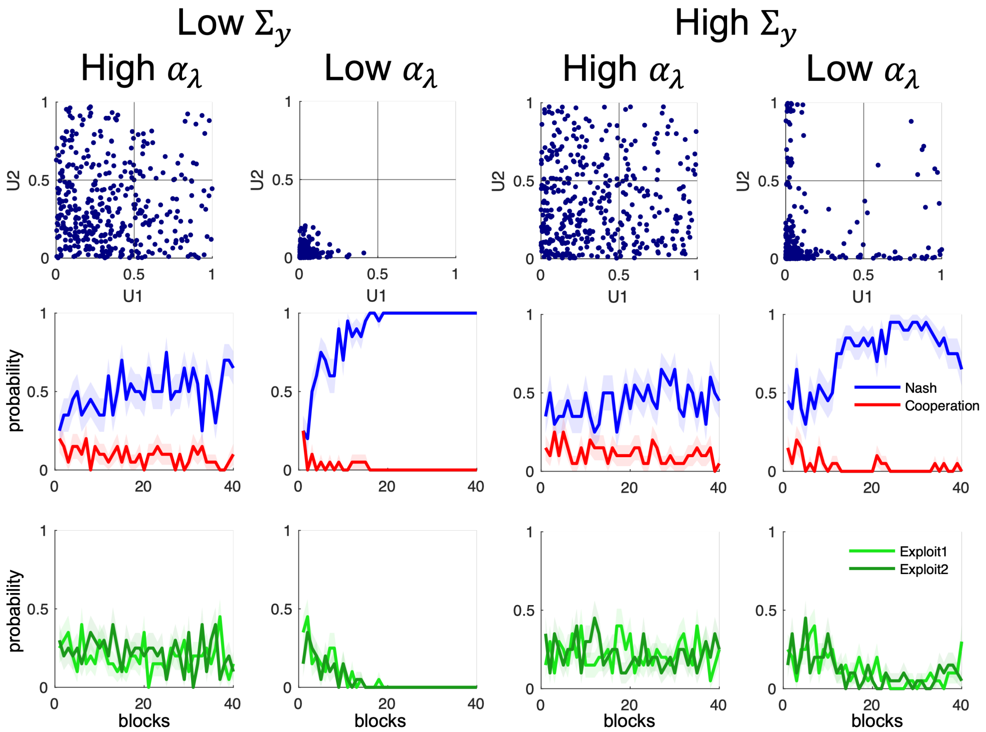

Figure 4.6 Prisoner dilemma simulation results. For each combination of parameters (Σ_y, λ and α_λ) we report data-points in the dyadic action space, probability over time to converge toward the Nash equilibrium, the cooperative solution and the exploit solutions.

4.4.2 Stag Hunt

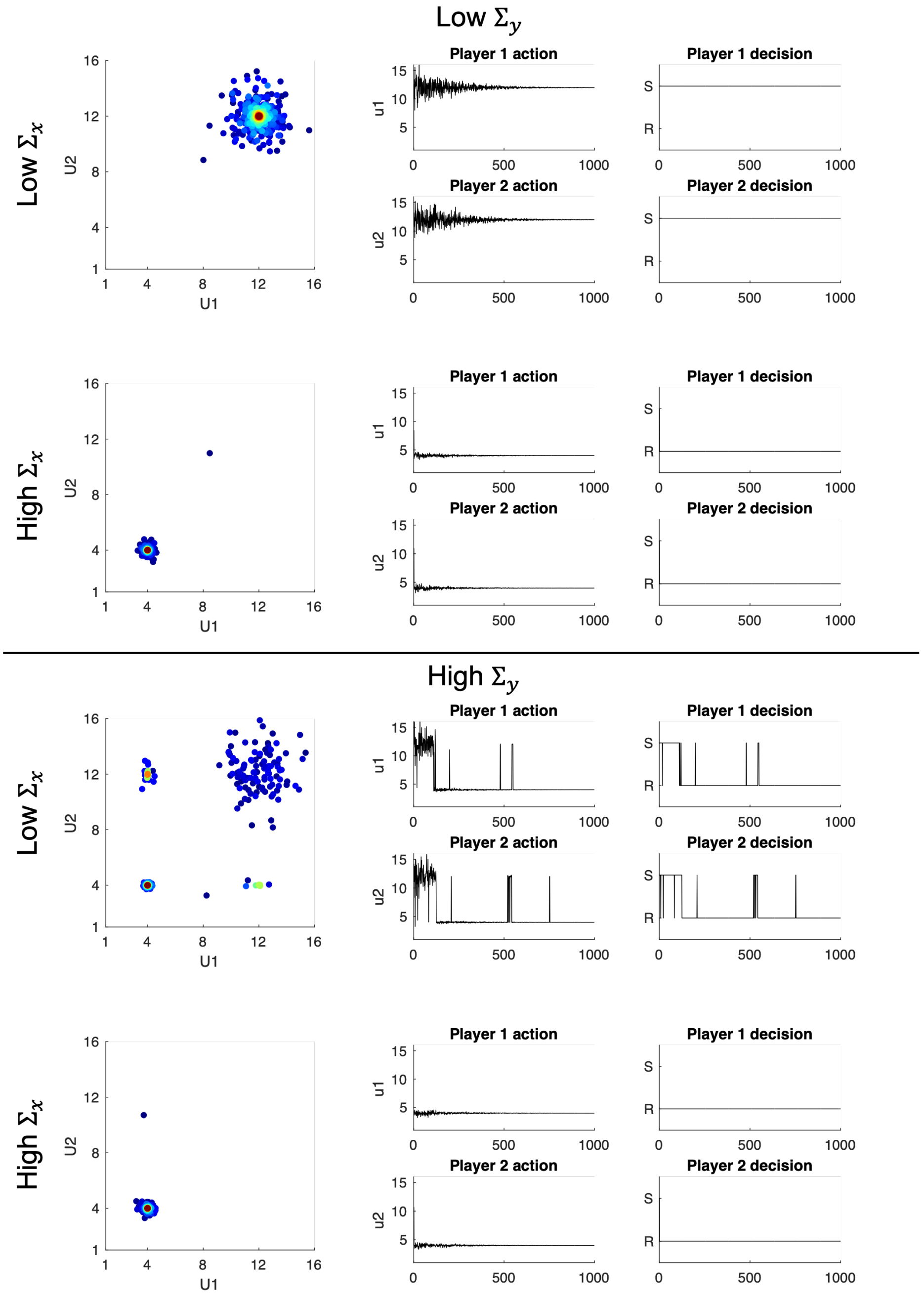

Figure 4.7 Stag Hunt simulation results. For each simulation, we report dyadic actions on the left. Datapoints color encode the trial number - cold colors represent initial trials, hot colors represent final trials. On the right we report Players actions and decisions (stag/rabbit). The first two line of plot were generated with low value of sensory noise, Σ_y and two different levels of internal noise, Σ_x.

Here, we report simulations results for a continuous version of the stag-hunt game, as proposed by Yoshida et al. (2008) - see Figure 4.7. We manipulated parameters in order to test the contribution of the partner model to the convergence toward one of the two possible Nash equilibria. In particular, here we report simulation results with two different values of sensory noise (Σ_y^i = 1 and Σ_y^i = 100) and two different values of internal noise (Σ_x^i = 1 and Σ_x^i = 100). When information about the partner is reliable (low sensory noise Σ_y) and players assume that their partner tends to have a predictable behavior from trial to trial (low), players tend to pursue the stag with increasing accuracy over time. With high internal noise, players are incapable to estimate partner actions and therefore developing a partner model. In this case, players tend to neglect the partner and decide to pursue the rabbit independently on the sensory system noise. When information about the partner is not reliable but players have good predictive capabilities, they tend to oscillate more between the different possible solutions but end up neglecting the partner, as it is the most rewarding solution. The results are consistent with the ones from Yoshida et al. (2008). Interestingly, it also provides a predictive tool that allows to make hypotheses that can be tested experimentally in sensorimotor versions of the same task.

4.4.3 2-VP task

Simulation We used the model as a generative tool to simulate two dyads, which only differed in the magnitude of the sensory variance, Σ_y: small ($\Sigma_y = 4 \cdot 10^{-4}$ m^2) and large ($\Sigma_y = 4 \cdot 10^{-5}$ m^2). As in the actual experiments, for each model we simulated a sequence of 120 trials. The simulation results are summarized in Figure 4.9, in terms of the time series of the minimum distances of Player i trajectory to the j-th via-point (MD$_{ij}$).

We observed that, consistently with experimental findings, with lower sensory noise players converged toward a shared solution which is close to the Nash equilibrium - figure 4.5.

Parameters identification We estimated the model parameters for each player and each dyad. As expected, and consistent with the experiment design, we found that the estimated sensory variance (Σ_y) decreases as the reliability of partner information increases (p = 0.0098) – see Figure 4.10. However, a significant group effect was also observed in the parameters which describe the action selection – greater λ_1 (p = 0.0308). Post-hoc analysis revealed a significant H-PV difference for measurement noise (p = 0.0072) and a significant H-VH for the initial exploration rate λ_1 (p = 0.0235).

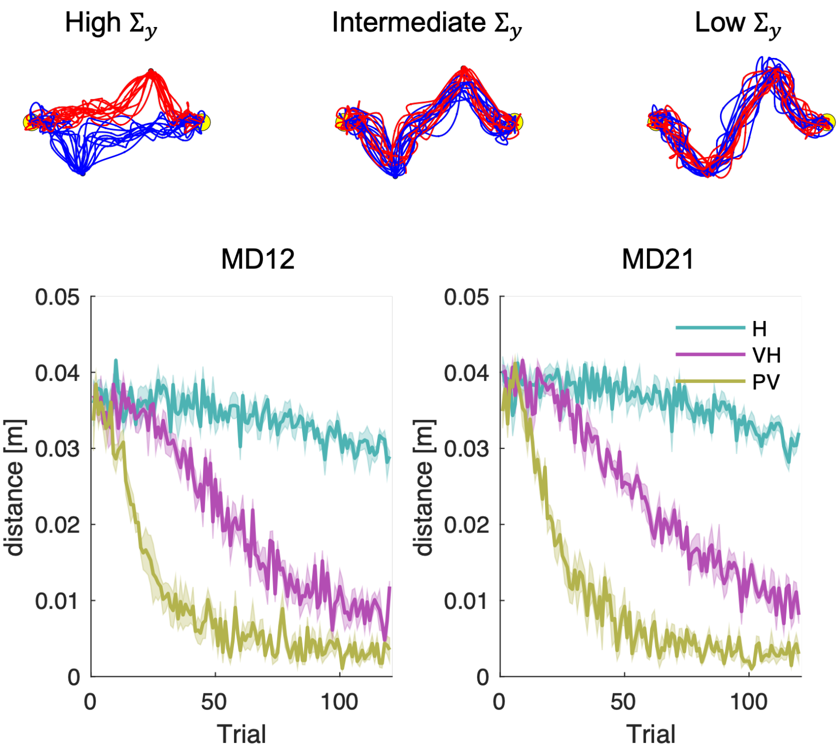

Figure 4.8 Simulated trajectories, for large (a) and small sensory noise (b). Top: movement trajectories. Bottom: minimum distances from self and other via-points, for Player 1 (blue) and Player 2 (red).

4.5 Discussion

We addressed the problem of how pairs of agents gradually negotiate a coordinated movement over repetitions of the same task. We specifically focused on how achieving a coordination is determined by the ability to predict partner actions. We presented a general computational framework to describe the mechanisms underlying the development of joint coordination. The model can be used to simulate joint action tasks and most importantly to analyze actual experiments. Previous models of joint action were indeed purely theoretic (Pesquita et al., 2018) or capable of simulating qualitatively interacting behavior (Chackochan and Sanguineti, 2019; Takagi et al., 2017). We also developed a fitting procedure which allows to quantitatively characterize experimental results in term of the proposed model.

We demonstrated the model in the context of three joint coordination scenarios: a continuous motor version of the classical Prisoner dilemma (Braun et al., 2009) and of the Stag-hunt games (Yoshida et al., 2008), and in the 2-VP task (Chackochan and Sanguineti, 2019).

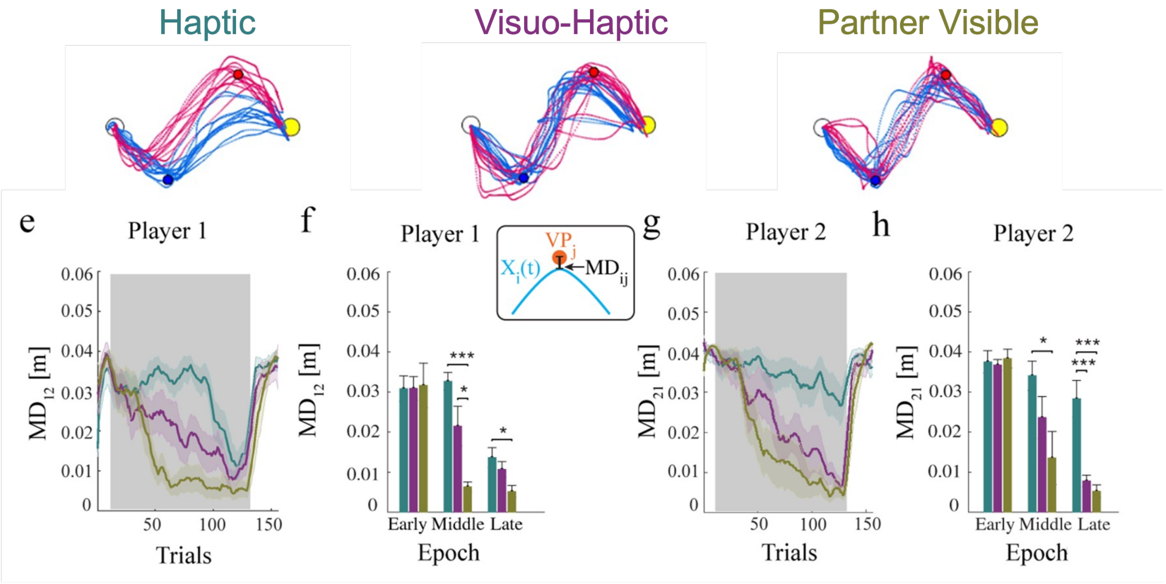

Figure 4.9 Trajectories for the three experimental groups tested in (Chackochan and Sanguineti, 2019) study. Top: movement trajectories. Bottom: minimum distances from other via-points for the two players in the three groups.

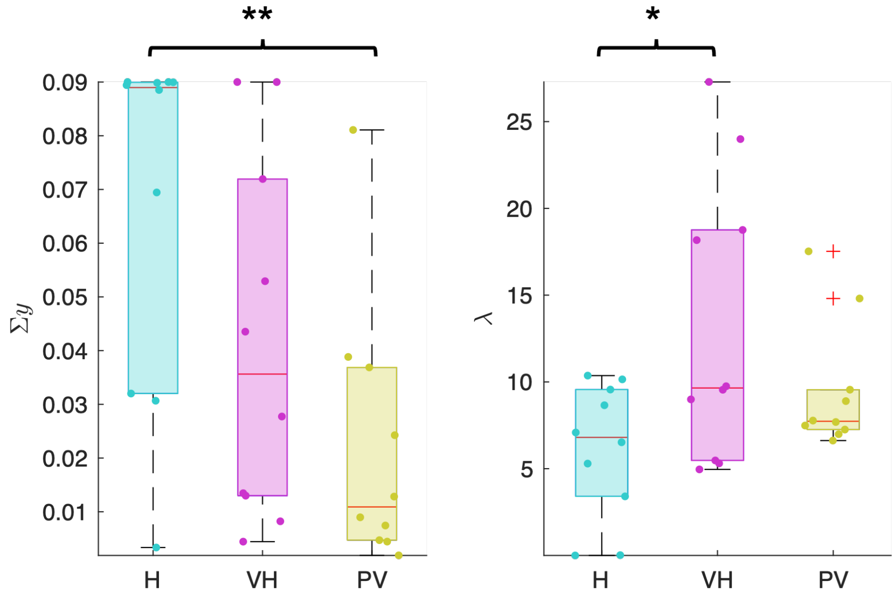

Figure 4.10 Group differences in the estimated model parameters. Boxplots display median, 25th and 75th percentiles of the estimated parameters values, for H (light blue), VH (pink) and PV (yellow) groups. Post-hoc analysis: $*$: $p < 0.05$; $**$: $p < 0.01$

In the Prisoner dilemma agents gather information about their partner through a variable force that prevents or facilitates their forward motions depending on their own and their partner's action. In the Stag-hunt, players gather the highest gain if they both decide to hunt a stag, while they gather smaller gains if they catch a rabbit. To pursue a stag they need to cooperate, while catching a rabbit does not require cooperation. In the 2-VP task pairs of agents are mechanically coupled, and have to perform reaching movements from the same start to the same endpoint by crossing a via-point while keeping the interaction force low. These games have one Nash equilibrium.

We modeled the joint tasks as a quadratic game and introduced uncertainty at both the perceptual and action selection level. Each player, privately observes the available signals and simultaneously chooses their action to minimize their cost function. In the Prisoner dilemma game, we were capable to generate dyadic behavior as observed in experimental settings (Braun et al., 2009). In particular, we show performance evolution over time. Players converge to the (defect-defect) solution as in the experiments (Braun et al., 2009). If the sensory variance is low and the temperature decay rate is low, players converge to the joint solution more quickly. Smaller values of temperature decay rate imply that players gradually become aware of their cost function, stabilizing their strategy. In the Stag-hunt game, we showed that the convergence toward the most rewarding solution (cooperative solution) requires reliable sensory information about the partner, but, most importantly, that players have the belief that the partner stably decides to catch a stag. As regards the 2VP task, experimental results suggested that information reliability affects the rate of convergence toward the equilibrium. For fitting purposes we rearranged the model, with each player i specified by a dynamical model where the input is the sequences of partner actions and the output is the sequence of the executed actions. The analysis of the experimental observations in terms of the model parameters suggests that (as expected from experimental design), the measurement noise Σ_y decreases from the H to the PV group, consistently with experimental conditions (Chackochan and Sanguineti, 2019). In addition, the greater initial temperature λ_1 of action selection observed in the VH condition suggests that displaying interaction force as a vector tends to trigger extra exploration, possibly because the information is not immediately interpretable during movements.

4.6 Conclusion

Overall, these findings confirm that model-based analysis of joint action experiments may provide valuable information on the underlying mechanisms. In particular, we addressed joint action using a couple of computational models representing the two interacting agents. The model can be used to address different interactive settings. We addressed a setting where the matrices defining the cost function are positive definite, then the conditional action probability is a Gaussian distribution. However, in competitive games the gain of one player corresponds to the loss of the other; in this setting, the cost functions may become negative, which implies that matrix R^i_{ii} may no longer be positive definite. In this case, the action distribution is no longer a Gaussian distribution, however, if the action space is limited the model is still valid. In conclusion, the model is very general and can be used to describe any static game with continuous action space. In the presented formulation we made a series of assumptions. We hypothesized that actions, partner model, and sensory information are continuous variables. We assumed that the partner model is a Markovian process, depending only on the previous step. Finally, we hypothesized the perfect knowledge of our own goal and no knowledge about our partner's goal. The model can be extended further, relaxing the above-mentioned assumptions. For instance, if players have no complete knowledge of their tasks, they may build a model of their own cost function which can be updated trial-by-trial based on a reward measure.

Coupled computational models that evolve over time to describe the structure and behavior of a physical system or process are referred to as digital twins. Digital twins and computational models of joint action are generally complicated frameworks which require custom implementations. The presented model relies on dynamical system theory and probabilistic graphical models. This is consistent with previous works (Kapteyn et al., 2021) proposed a general and robust foundation for the development of digital twins. As such, the model described in this chapter is a promising tool to advance the understanding of the processes underlying joint action and a wide variety of applications can be envisioned. In healthcare, computational models or digital twins of interacting agents promise to advance medical assessment, diagnosis, and personalized treatment (Bruynseels et al., 2018; Rivera et al., 2019). In a similar way, in educational settings, they can be used to devise personalized training paths.

Chapter 5

Learning and partner representation in joint coordination

5.1 Introduction

Two persons walking one toward the other in a narrow hallway while avoiding bumping one into the other is just one example of the kind of interactions that we deal with every day. Scenarios like this typically involve a combination of selecting one among a set of possible actions – for instance, keep left or keep right – and sensorimotor coordination to adjust the spatiotemporal features of the movement. Action selection may be the result of a shared decision, involving explicit communication of goals and/or intentions among the participants (Vesper et al., 2017). Or, each individual may select their action independently, on the sole basis of their own expected payoff. The latter likely accounts for the information available about their counterpart, but the extent to which this information is represented and/or accounted for in action selection is a matter of debate (Chackochan and Sanguineti, 2019; Kourtis et al., 2013; Loehr and Vesper, 2016; Lokesh et al., 2022; Sebanz et al., 2005; Takagi et al., 2017; Vesper et al., 2017; Yamaguchi et al., 2017).

Some authors suggested that interacting agents represent the partner's ongoing action (Chackochan and Sanguineti, 2019; Gillijns and De Moor, 2007; Lokesh et al., 2022). Others suggested that players may represent partner goal, i.e cost function (Li et al., 2019; Takagi et al., 2017), or intentions (Sebanz et al., 2005). While some studies proposed that in joint action humans develop only representations at the individual level (Yamaguchi et al., 2017), it has also been stressed the fact that in joint action humans not only represent individual contributions but also develop representation at the group level (Kourtis et al., 2013; Loehr and Vesper, 2016).

As introduced in previous chapters, joint action can be interpreted and analyzed in terms of game theory. Interaction tasks can be modeled as games. Several studies have reported that the selected actions converge to Nash equilibria (Braun et al., 2009, 2011; Chackochan and Sanguineti, 2019; Grau-Moya et al., 2013; Leibfried et al., 2015). If there is uncertainty in the action selection process, joint actions converge to quantal response equilibria (Lindig-León et al., 2021a; McKelvey and Palfrey, 1995, 1998).

The way we converge to an equilibrium - i.e learn coordination - is still another issue. If own payoff function is not completely known, reinforcement learning may provide a suitable framework to address these scenarios. The gradual evolution and achievement toward the equilibrium have been recently modeled with Q-learning (Lindig-León et al., 2021b). Other studies proposed fictive play (Fudenberg and Levine, 1998) as a possible learning rule in joint action (Chackochan and Sanguineti, 2019; Grau-Moya et al., 2013). Fictitious Play is a possible general learning mechanism (Fudenberg and Levine, 1998) where interactive agents chose their best next action based on an internal representation of the partner.

A previous study (Chackochan and Sanguineti, 2019) focused on a situation in which participants are mechanically coupled, have a shared final target and are instructed to cross a via-point (VP, different for each player) while keeping the interaction force at a minimum. Hereafter this experiment will be referred to as the joint two via-point (2-VP) task. In terms of game theory, the joint 2-VP task is a pure coordination game with two Nash equilibria that are very different in terms of cost. The low-cost Nash equilibrium corresponds to both players moving through both VPs in the same temporal order (VP1 first or VP2 first, whatever leads to the shortest path). Over repeated iterations of the game, the two players converged to a stable coordination. It has been suggested that convergence is achieved by assuming that at each movement repetition (trial) each player estimates the action of their opponent (partner model), and determines their next action through a form of fictitious play. This simple strategy automatically accounts for the amount of information available about the partner.

Here we present a particular case of the 2-VP task where VPs are arranged symmetrically - i.e. both VPs are equidistant and symmetric with respect to the start-target line. This game has two cost-equivalent equilibria. To develop stable coordination, the players need to converge together to either one or the other. Using this task, we not only address the convergence toward an equilibrium (shared solution) but we investigate the processes underlying equilibrium selection. We hypothesized that more reliable information about the partner would lead to more stable coordination. Similar to Chackochan and Sanguineti (2019), we tested two dyads' groups in which we manipulated the information available about the partner. We first characterize the decision-making/action selection problem empirically. We then used the computational model proposed in Chapter 4 to address scenarios in which there are multiple equilibria. We estimate model parameters from the experimental data and show that model captures the different observed behaviors.

5.2 Materials and methods

5.2.1 Experimental Apparatus and Task

Each experiment involved two participants (a dyad). Each participant sat in front of a computer screen and grasped the handle of a haptic interface (Novint Falcon). They could not see or hear each other and were not allowed to talk. The experimental apparatus is depicted in Figure 5.1A.

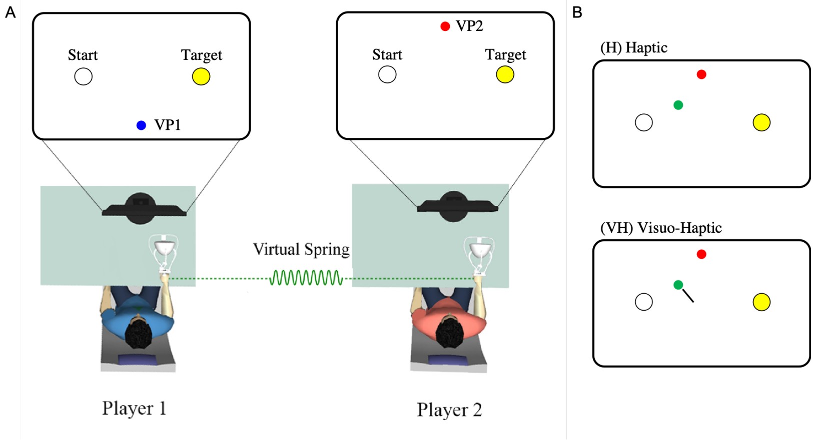

Figure 5.1 Experimental apparatus and protocol. A: Partners in a dyad were connected through a virtual spring. Players in a dyad were connected through a virtual spring. Both players were instructed to perform reaching movements in the vertical plane, between the same start point and the same target point but through different via-points. Each player could only see their own VP, but not their partner's. Both were instructed to keep the interaction force as low as possible during movement. The experimental protocol consisted of three phases: baseline, training, and wash-out. During the baseline phase, the interaction forces were turned off, and each player performed on their own ('solo' performance). The players were mechanically connected during the training phase, and the connection was permanently removed during the wash-out phase. B: As in (Chackochan and Sanguineti, 2019), we manipulated the information available on the partner's actions by only providing it haptically, through the interaction force (Haptic group, H) or by additionally displaying the interaction force vector on the screen (Visuo-Haptic group, VH)

The participants were instructed to perform planar point-to-point movements in the vertical plane, between the same start and end point (displayed, respectively as a white and a yellow circle of $\oslash$ 1 cm), but through different via-points (a blue circle for Player 1, a red circle for Player 2, both with $\oslash$ 0.5 cm. In a reference frame centered on the robot workspace, with the X axis aligned with the left-right direction and the Y axis aligned with the vertical direction, for both players the start point was placed in the (-5, 0, 0) cm position and the target point was placed in the (5, 0, 0) cm position. Hence the start and the target point had a horizontal distance of 10 cm. The via-points were different for the two participants and were placed, respectively, at locations VP1 = (0,-3,0) cm and VP2 = (0,3,0) cm, symmetric with respect to the Y axis. A trial started when both participants placed their cursor inside the start region. Then the target and a via-point ($\oslash$ 0.5 cm circle) appeared. The participants were also instructed to keep their movements as planar as possible, i.e. by keeping the depth, Z coordinate within the range (±4 cm) from the origin of the workspace. The current positions of the end effectors, x_1 and x_2, were continuously displayed to each partner, as $\oslash$ 0.5 cm circular cursors on their respective screens, colored in green if the depth was correct and in red otherwise.

The experimental protocol was organized into epochs of 12 movements each and consisted of three phases: (i) baseline (1 epoch), (ii) interaction (17 epochs), and (iii) wash-out (2 epochs) for a total of 20 × 12 = 240 movements. During the baseline phase, the interaction forces were turned off and each subject performed on their own ('solo' performance). During the interaction phase, the two participants were mechanically connected through the haptic interfaces, which generated a force proportional to the difference of the two hand positions, i.e. $F_i = -k(x_i - x_{-i})$, with k=300 N/m. In randomly selected trials (catch trials) within each epoch (1/6 of the total, i.e. 2 trials per epoch) the connection was removed. The connection was permanently removed during the wash-out phase. The total duration of the experiment was about 45 minutes. The participants were told that they might experience a force while performing the task and were instructed to keep this force to a minimum. We introduced this additional requirement because of a technical limitation of our robots, which are unable to generate large stiffnesses. With a low stiffness the haptic perception of the interaction force is less reliable, and introducing an explicit requirement on low interaction forces made sure that all players were provided with exactly the same task requirements irrespective of the information they had available about their partner. At the end of each movement, each subject received a 0-100 reward, calculated as a function of the minimum distance of their movement path from their own via-point and of the average interaction force: $\text{score}_i = \frac{100}{1+\exp^{h(d_i - d_0)}}$, where $d_i = d_{\text{VP}_i} + c \cdot d_{12}$ and $i = 1,2$. The quantities d_{VP_i} and d_{12} are, respectively, the minimum distance between the movement trajectory and the subject's own 'via-point' (VP_i) and the average distance between the two participants' hand positions. In the disconnected trials, we took $c = 0$, i.e. the score only depended on how close the participants got to their own via-point. Parameters h and d_0 were calculated so that the score was maximum (100) for $d_i \leq 0.005$ m (i.e., the VP radius), and minimum (0) for $d_i \geq$ 0.02 m. To encourage participants to establish a collaboration, in trials in which the two participants were mechanically connected we took $c = 0.5$, so that to get a maximum score participants also had to keep their relative distance as low as possible. The score is only meant as a reinforcement signal to sustain players' motivation and to speed up learning. As such, it specifically focuses on a few aspects of the task (pass through the via-point; keep the interaction forces low). Other task requirements, like reaching the target, required no reinforcement. To encourage participants to maintain an approximately constant movement duration, after each movement, a text message on the screen and changes in the color of the target (either green or red) warned the participants if the movement was either too fast (duration < 1.85 s) or too slow (duration > 2.15 s). However, the participants received no penalization if their movement duration did not remain within the recommended range.

Each dyad was randomly assigned to two groups, which differed in the available sensory information. In the haptic (H) group, the interaction could only be sensed haptically. In the visuo-haptic group (VH), the interaction force (magnitude, direction) was also displayed as an arrow attached to the cursor (scale factor: 10 N/cm). Hence the participants in the VH group had more reliable information about partner's movement – see Figure 5.1B.

5.2.2 Subjects

The 36 study participants were selected among the graduate and undergraduate students of the University of Genoa. Right-handedness was assessed for all participants by using the Edinburgh Handedness Inventory (Oldfield, 1971). All participants were naive to the task and had no known history of neurological or upper limb motor impairment. Participants were paired by matching their body mass index (BMI), thus forming 18 dyads. The two participants within the same dyad were randomly labeled as, respectively, Player 1 and Player 2. We assigned each dyad to two groups: H (25 ± 3 y; 6 M +12 F) and VH (24 ± 5 y; 8 M + 10 F) groups. The study conforms to the ethical standards laid down in the 1964 Declaration of Helsinki that protects research subjects and was approved by the Institutional Review Board (Comitato Etico Regione Liguria). All participants provided written informed consent to participate in the study.

5.2.3 Data analysis

During experiments, hand trajectories and the interaction forces generated by the robots were sampled at 100 Hz and stored for subsequent analysis. Movement and force trajectories were smoothed using a 4th order Savitzky-Golay filter with a 370 ms time window. The same filter was used to estimate velocity and acceleration. For every trial and each player within a dyad, we identified the start and end times of each trajectory, as the time instants ST_i and ET_i at which the speed of player i crossed a 2 cm/s threshold. At dyad level, we took $\mathrm{ST} = \min\left[\mathrm{ST}_1, \mathrm{ST}_2\right]$ and $\mathrm{ET} = \max\left[\mathrm{ET}_1, \mathrm{ET}_2\right]$ as the dyad start and end movement times.

Coordination

We characterized coordination in terms of both movement kinematics and kinetics, at both individual and dyad level. As regards individual players, coordination can be understood in terms of game theory. The symmetric 2-VP game has two Nash equilibria, in which both players pass through both via-points by following the same trajectory, either in $\mathrm{VP}_1 \rightarrow \mathrm{VP}_2$ or $\mathrm{VP}_2 \rightarrow \mathrm{VP}_1$ order; see Figure 5.2A. This prediction suggests that whatever the order is, a sign of coordination is that each player, while passing through their own via-point, also gets very close to their partner's Chackochan and Sanguineti (2019). This can be quantified in terms of the Minimum Via-Point Distance, MD_{ij} of player i from via-point j: $\mathrm{MD}_{ij} = \min_t \mathrm{dist}_{ij}(t)$, with $\mathrm{dist}_{ij}(t) = ||x_i(t) - x_{\mathrm{VP}_j}||$ and $i, j = 1, 2$.

At dyad level, coordination implies a reduced interaction force between the players. We calculated the average interaction force $\mathrm{IF} = \frac{1}{N} \sum_t \|F(t)\|$, where $F(t)$ is equal and opposite for the two partners in the dyad.

Another crucial indicator of stable coordination at dyad level is the temporal synchronisation of players' movements. To quantify this, we looked at the difference in start times of the two players within a dyad, $\Delta\mathrm{ST} = |\mathrm{ST}_1 - \mathrm{ST}_2|$. Low values of this quantity indicate that players start to move together, on the contrary, higher magnitude indicates that players are not coordinating their actions.

We also looked at the Crossing Time (TC_{ij}), defined as the time instant – with respect to the dyad start time – at which player i achieves such a minimum distance to via-point j: $\mathrm{TC}_{ij} = \arg\min_t \mathrm{dist}_{ij}(t)$; therefore, $\mathrm{MD}_{ij} = \mathrm{dist}_{ij}(\mathrm{TC}_{ij})$. In case the players coordinate, the crossing times of the two players when passing through the same via-point would tend to be as close as possible. To quantify synchronization we observed the difference of crossing time to each via-point i of the two players within a dyad, $\Delta\mathrm{TC}_i = |\mathrm{TC}_{1i} - \mathrm{TC}_{2i}|$. Where lower values indicate that players cross the via-point with the same timing.

Action selection

A pre-requisite for coordination is that both players cross their partner's via-point, i.e. both reduce their MD_{ij} with $i \neq j$. We classified each trial as collaborative in space if both players had $\mathrm{MD}_{ij} < d$, with $d = 0.02$ m. The threshold was set based on the observation of dyads who clearly exhibited a coordination. An additional requirement for coordination is that players within a dyad cross the via-points in the same order. This is challenging, as there are two equally effortful trajectories. To assess this, we looked at temporal coordination. For each trial in which both players reduced their MD_{ij} below the threshold, each player could either cross VP1 first and then move toward VP2 (VP1→ VP2 or 1-2 strategy) or cross VP2 first and then VP1 (VP2→VP1 or 2-1 strategy). These strategies can be identified by looking at the difference between the crossing times of a given player, $\Delta\mathrm{TC}_i = \mathrm{TC}_{i1} - \mathrm{TC}_{i2}$. In other words, we considered the temporal order in which the minimum distances from own and partner via-points were reached (own via-point) or approached (likely, partner's via-point). We assessed whether player i crossed/approached VP1 or VP2 early or late in the movement. Based on this analysis, for each player we defined a decision time series in which at every trial t each player, $i = 1,2$, selects his/her action $s_i(t)$ among three options: moving to VP1 and then VP2 (1-2); moving to VP2 and then VP1 (2-1); or moving to own VP while ignoring the other (1-1 or 2-2). In summary:

$$s_i(t) = \begin{cases} \text{1-2} & \text{if } (\mathrm{MD}_{i1} < d) \wedge (\mathrm{MD}_{i2} < d) \wedge (\mathrm{TC}_{i1} < \mathrm{TC}_{i2}) \\ \text{2-1} & \text{if } (\mathrm{MD}_{i1} < d) \wedge (\mathrm{MD}_{i2} < d) \wedge (\mathrm{TC}_{i1} > \mathrm{TC}_{i2}) \\ \text{1-1} & \text{if } \mathrm{MD}_{i2} > d \\ \text{2-2} & \text{if } \mathrm{MD}_{i1} > d \end{cases} \tag{5.1}$$

We then compared the time series describing the players' actions. We classified each trial as collaborative, at dyad level, if both players used either 1-2 or 2-1:

$$\mathrm{collaboration}(t) = [(s_1(t) = \text{1-2}) \wedge (s_2(t) = \text{1-2})] \vee [(s_1(t) = \text{2-1}) \wedge (s_2(t) = \text{2-1})] \tag{5.2}$$

Using this definition, for each dyad we calculated the temporal evolution of the probability of collaboration, Pr(collaboration) over epochs. In order to understand the determinants of learning, we defined a player as a 'learner' if he/she consistently achieved either 1-2 or 2-1 during the last epoch of the training phase; 'non-learner' otherwise. Note that both players in a dyad being learners according to this definition does not necessarily imply that they achieve a collaboration – they may get stuck into a cyclic behavior in which they alternate 1-2 and 2-1 strategies and never converge.

Model

To help interpreting the empirical observations in terms of the mechanisms underlying the development of a joint coordination, we used the model described in the previous chapter 4. We modeled the action selection and spatiotemporal coordination process as a static non-cooperative infinite game (Başar and Olsder, 1998) with incomplete information. In particular, each player (i) knows their own goal, but not their partner's; and (ii) has limited (uncertain) information about partner's action. We also assume that (i) players predict partner actions by optimally combining sensory and prior information; and (ii) learning of a coordination strategy occurs through fictitious play (FP) (Fudenberg and Levine, 1998).

The model addresses the trial-by-trial evolution of players' action and assumes that no adaptation occurs in the ongoing movements within a single trial. In other words, the model assumes that an action within a singe trial mostly develops feedfoward. This is a reasonable assumption given the short duration of one individual trial. At each trial t, the model assumes that i-th player selects an action, $u_i(t)$. Joint action scenarios like the symmetric 2-VP joint task, have multiple strategy options. To express this, a categorical variable, $s_i \in \left\{S_i^1 \cdots S_i^M\right\}$, specifies one of the M strategies to be selected. Both strategy choice and action selection are assumed to depend on the partner's expected move (partner model), $x_i(t) \equiv \hat{u}_{-i}(t)$. Once the action has been carried out, through their sensory system the player collects information $y_i(t)$ about their partner's action. This information is then used to update the partner model, which will be accounted for when selecting the action on the next trial. Figure 5.2 summarizes the model architecture for one individual player, but both players at the same time undergo this process. Therefore, the whole joint action is described by a pair of these models – one per player.

We used the model to identify model parameters for each player in a dyad from the time course of dyad movements. From the identification procedure, we obtained a set of parameters for each player. Parameters constitute additional metrics that characterize interacting agents in terms of their capabilities to gather information about the partner (Σ_y), to integrate such information into a partner model (A, Σ_x, $\hat{P}_0$), to use the partner model to select an appropriate action (λ_1, α).

Statistical analysis

We expect that coordination and decision-making evolve with time (learning). We also asked if they are affected by the amount of information each player has available about their partner. As for coordination, we ran a repeated-measures ANOVA with group (H, VH) and Time

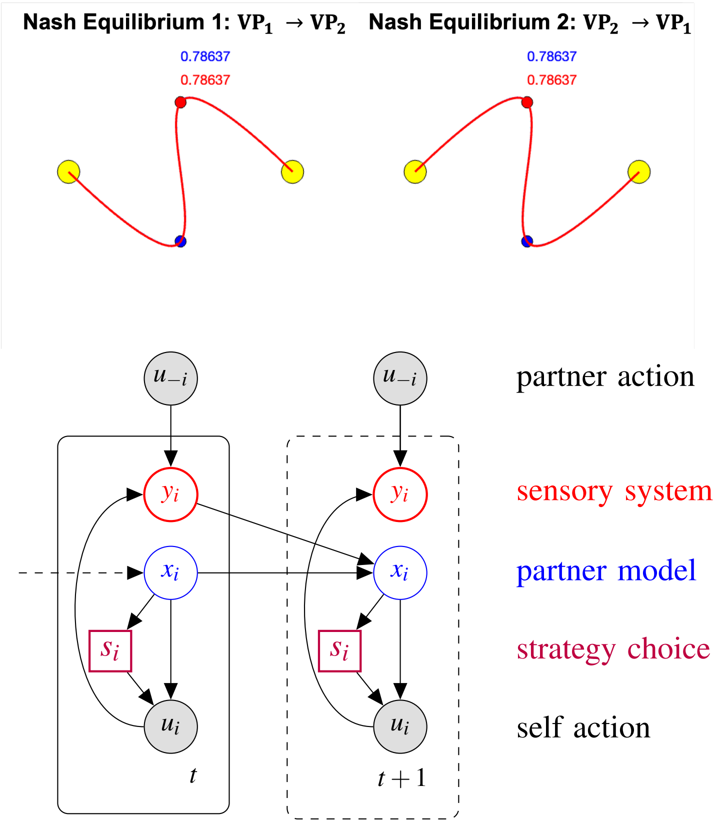

Figure 5.2 In the upper part of the panel the two Nash equilibria are reported for this 2 via-points task. In the lower part of the panel the computational model is outlined. The computational model is extended to account for multiple strategy choices. At each time step, Player i receives information about partner action (u_{-i}) through their sensory system (y_i). This information is used to predict the next partner action (partner model, x_i). The partner model is used to select the actual action (u_i) which also depends on the strategy choice (s_i). The latter decision is also affected by the player's knowledge about the partner. For each player, sensory input (y_i), partner model (x_i) and strategy choice (s_i) are latent variables; actions (u_i and u_{-i}) are observable variables (denoted by shaded nodes).

(early – training epoch 1 – and late – epoch 17) as factors, for minimum distance (MD_{ij}), crossing times (TC_{ij}) and interaction forces (IF). As regards action selection, we focused on the selection of coordinated actions (either 1-2 or 2-1) over non-collaboration (1-1 or 2-2) and, at dyad level, on collaboration versus non-collaboration. We specifically looked at how $\Pr(u_i(t) = \text{1-2} \vee u_i(t) = \text{2-1})$ and $\Pr(\text{collaboration}(t))$ evolve with training. We specifically assessed whether (i) $\Pr(\text{collaboration}(t))$ evolves with time; and (ii) $\Pr(\text{collaboration}(t))$ exhibits any group effects. Because of the probabilistic nature of the data to be compared, we used the non-parametric Wilcoxon Rank-Sum Test. We used a one-way ANOVA to analyze group effects (H vs VH) for each of the estimated model parameters. We used the same test with group Learner/Non Learner as a factor. Further, we used a Linear Discriminant Analysis to understand whether estimated parameters can be used to classify our players into H/VH and Learner/Non Learner, as described in the previous section 5.2.3.

All data analysis and simulations were carried out using custom software written in MATLAB (R2022a). Statistical significance was considered at $P < 0.05$ level for all tests.

5.3 Results

5.3.1 Experiments

The symmetric version of the 2-VP task is more challenging than its asymmetric counterpart (Chackochan and Sanguineti, 2019) as there are two equally costly coordination strategies (Nash equilibria) – see Figure 5.2A. By looking at the participants' trajectories and their evolution over trials, we observed three main behaviors. In some dyads, players just ignore each other and only focus on their own via-point. In other dyads, both players converge to a common strategy (either 12 or 21). Other dyads exhibit a cyclic behavior, in which on each trial the players choose different strategies (either 12 and 21, or 21 and 12) and they both switch strategies on the subsequent trial. Figure 5.3A exemplifies these behavior types.

Coordination

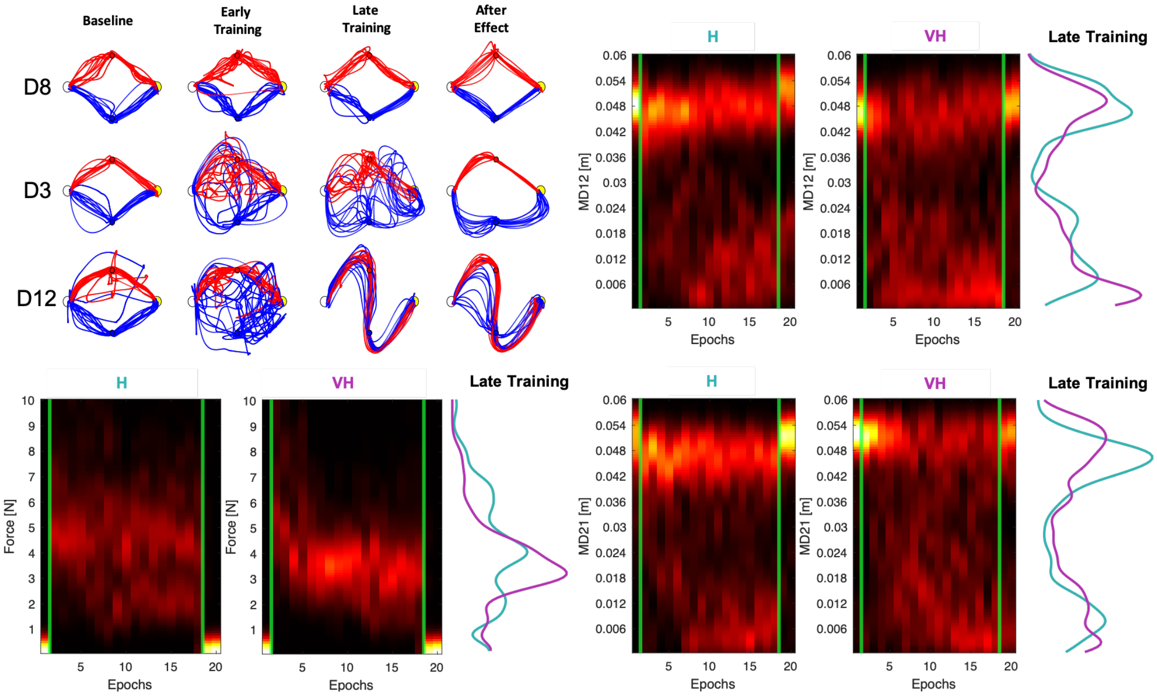

Figure 5.3 Dyads exhibit a variety of coordination behaviors. A: Typical dyad trajectories (from top to bottom: D8, D3, D12) at baseline, early/late training, and washout phases of the experiment. Player 1 and Player 2 are depicted in blue and red. These dyads exemplify typical behaviors: players ignoring each other (D8); cycling between opposite strategies (D3); and converging to a coordination (D12). B: Distribution of minimum distance to partner via-point, for Player 1 (MD$_{12}$, top) and Player 2 (MD$_{21}$, bottom), for dyads in the H (left) and VH group (right). C: Distribution of interaction force (IF), for dyads in the H (left) and VH group (right). The curves on the right display the probability density profiles of MD$_{12}$, MD$_{21}$, IF at end training (epoch 18,) for the H and VH groups (respectively, in magenta and green). A fraction of the dyads has learned to coordinate (low MD$_{ij}$ and IF), other dyads have not. For visualization purposes, the distributions have been approximated with a kernel smoothing function (bandwidth: 2.5 mm and 0.4 N)

The overall coordination performance is summarised in Figure 5.3B,C, D. In some dyads, one or both players gradually reduce their minimum distance from the opponent via-point – see Figure 5.3C. The interaction force also decreases with training; see Figure 5.3B. In other dyads, neither minimum distance nor interaction force reduce over training.

Statistical analysis confirms an overall significant Time effect for both players (MD$_{12}$: $p = 0.0235$, $F(1,16) = 6.272$; MD$_{21}$: $p = 0.0069$, $F(1,16) = 9.621$), but neither a significant Group effect, nor a significant Time × Group interaction.

Likewise, the interaction force exhibited no significant Group effect nor a significant Time × Group interaction, but we found a strong Time effect – the interaction force gradually decreases with trials (p=0.0002, F(1,16) = 23.187).

We looked at the distribution of the MD_{ij} and observed a bimodal trend at the end of training (epoch 18), indicating that for both experimental conditions there is a fraction of players that understand and attempt to incorporate partner's action in their motor command. However, reducing the minimum distance from the partner via-point does not necessarily result in stable joint coordination. From this observation, we classified a player as a learner if their average MD_{ij} was below a threshold (0.02 m) at the end of training. We found that 7 players out of 18 were classified as Learner in the Haptic group and 9 players out of 18 in the Visuo-Haptic group. However, the fraction of Learner over the total exhibited no significant group difference χ^2.

We performed a chi-square test to determine the two proportions' statistical difference, and we found no difference.

In summary, all measures of coordination improve with time even though only a fraction of dyads exhibit a stable decrease of MD and/or IF. Further, being a Learner does not seem related to the amount of information about the partner. Hence kinematic and kinetic measures are not sufficient to describe the coordination mechanisms and the determinants of convergence to stable coordination.

Stable coordination also implies that both players within a dyad synchronize the timing of their movements, in particular the time at which their movements start and end, and the time at which players get close to both via-points.To address coordination we looked at the difference of via-point crossing time between the two players - i.e. $\Delta\mathrm{TC}_i = |\mathrm{TC}_{1i} - \mathrm{TC}_{2i}|$.

Coordination is achieved if both players start their motion, cross the same via-point and end their motions with the same timing. We found a significant effect of Time on synchronization at the start and end point (Start: p = 0.0397, F(1,16) = 5.012; End: p = 0.0116, F(1,16) = 8.117) but we found no Group nor interaction Time × Group.

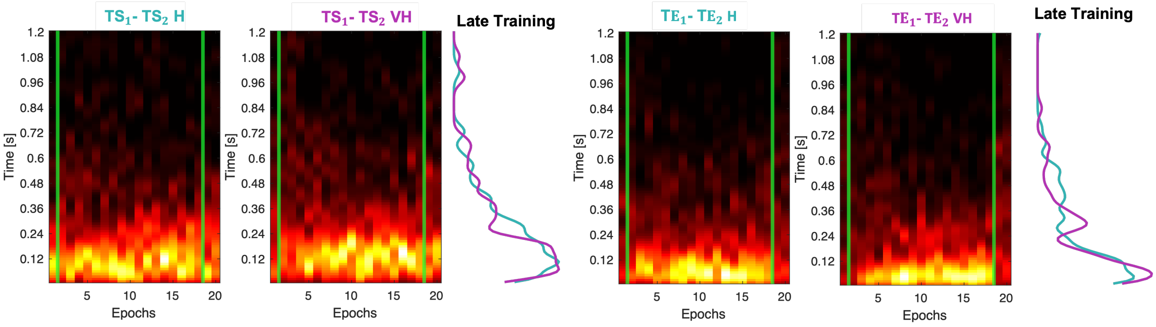

Figure 5.4 Initial and Final Synchronization.

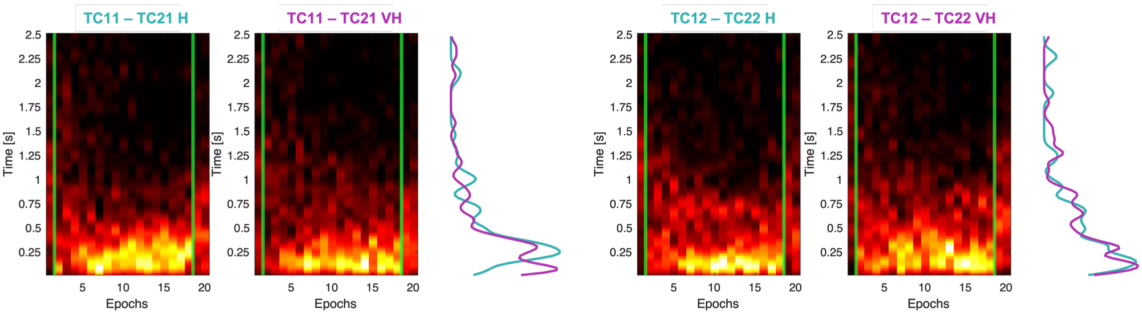

Figure 5.5 Synchronization at via-points.

When looking at synchronisation measure at VP_1 and VP_2 again we found a significant Time effect (ΔTC$_1$: p = 0.0065, F(1,16) = 9.783; ΔTC$_2$: p = 0.0058, F(1,16) = 10.128), but no Group effect or Time × Group interaction.

Overall, these findings suggest that increasing the reliability of information about the interaction does not have an effect on performance. Hence forms of coordination which require the selection of one among multiple cost-equivalent solutions is a much more complex process with respect to just converging to a unique equilibrium as in the asymmetric 2-VP setting (Chackochan and Sanguineti, 2019). However, a fraction of the dyads succeeds in achieving stable coordination in both space and time. To understand the underlying mechanisms, we focused on the decision-making component of the task i.e. whether each player decides to cross their VP earlier or later in the movement. We characterized the actions in space and time and carried out a probabilistic analysis to confirm the previous observation.

Action selection

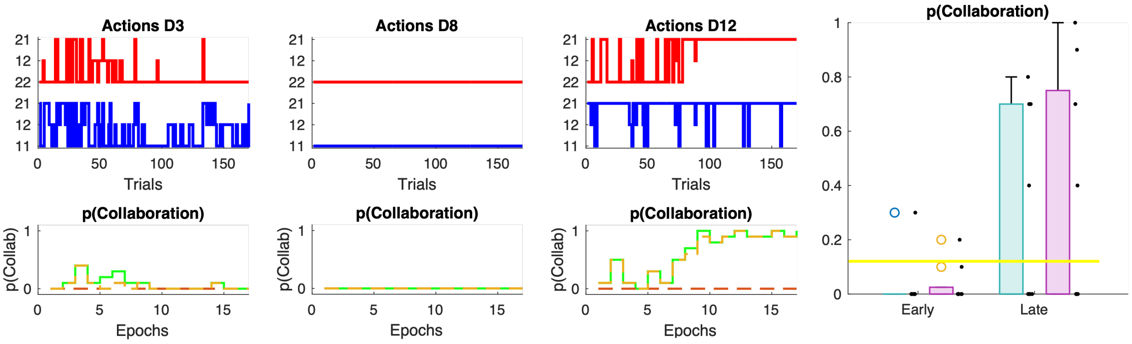

Figure 5.6 Dyads exhibit different action selection strategies. Data are reported for three representative dyads (D8, D3, D12). For each dyad, in the upper plot, we reported the individuals actions (Player 1: blue; Player 2; red). In the bottom plot we reported the probability to collaborate in green, the probability to collaborate using strategy 1-2 in dark orange and the probability to collaborate using 2-1 in yellow. On the right we report the boxplot of probability to collaborate in the early andlate training for the two groups in green (H) and purple (VH).

Looking at the trajectories, we observed that irrespective of the experimental group, individual participants adopted different behaviors. In particular - see Figure 5.6, we observed that players can decide to ignore the partner or to coordinate with them by attempting to cross also their via-point. However, at dyad level, this leads to four different behaviours: (i) both players ignore each other, (ii) one player ignores the partner but the other is more compliant and decide to cross also partner via-point, (iii) both players decide to collaborate and converge toward a shared solution and (iv) both players decide to collaborate but they never converge toward a shared solution. To check whether there are Time and Group effect we compared the probability of collaboration in the initial and final training epoch for the two groups, haptic and visuo-haptic. We found a significant Time effect ($p = 0.0064$, $F(1,32) = 8.52$) but no Group effect, consistent with results presented in the previous section. These results well described observed behaviors and confirm that in this scenario where two effort-equivalent solutions exist, it is harder to achieve stable coordination and the latter does not seem to depend much on the reliability of information about the partner.

Model

For each player within each dyad, we estimated the model parameters from the movements time series.

Consistent with the experimental findings, no parameter exhibited an H vs VH group difference. We then examined differences in the Learner vs Non-Learner groups. In this case, two model parameters exhibited a Group effect, namely the sensory variance (Σ_y, smaller in the Learner group, p = 0.0002, F(1,34) = 16.89) and the initial action selection temperature ($\lambda 1$, greater in the Learner group; p = 0.0062, F(1,34) = 8.53).

However, these tentative conclusions rely on analyzing one parameter at a time. In order to provide a multivariate analysis of the parameter space, we carried out a Linear Discriminant Analysis for both the H-VH and the Learner-Non Learner classes. In this way, we determined the unit vector w which maximally separates the two classes when the parameter vector x is projected as $z = w \cdot x$. We then quantified the discriminant ability of the parameter set x to discriminate among these groups in two ways. We first looked at the performance (area under the ROC curve, AUC, sensitivity, and specificity) of a scalar Naive Bayes classifier relying on the projected z. We then focused on the resulting discriminant vector w. The magnitude of each w component reflects the relevance of the corresponding model parameter in discriminating between the two groups. Consistent with the previous results, we found that the model parameters as a whole are poor predictors of the experimental groups H/VH (AUC = 0.64815; sensitivity = 0.5556; specificity = 0.7778 and F1 score = 0.7000). In contrast, model parameters clearly distinguish between Learners and Non-Learners (AUC = 0.8875, sensitivity = 0.85, specificity = 0.6875, and F1 score = 0.8095). For visualization purposes, we showed the data points (one per subject) in the principal components space. We observed that the two groups Learner and Non-Learner are quite clustered - see Figure 5.8. The magnitudes of the individual w components confirm the importance of sensory uncertainty (smaller in learners) and the initial action selection temperature (greater in learners). Learners also have greater partner model retention and smaller partner model randomness.

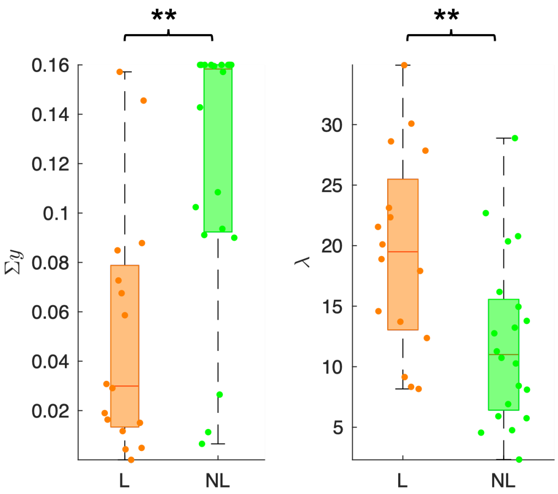

Figure 5.7 Estimated model parameters grouped in terms of Learner and Non-Learner. Boxplots display median, 25th and 75th percentiles of the estimated parameters values, for Learners (orange) and Non-Learner (green). $*$: $p < 0.05$; $**$: $p < 0.01$.

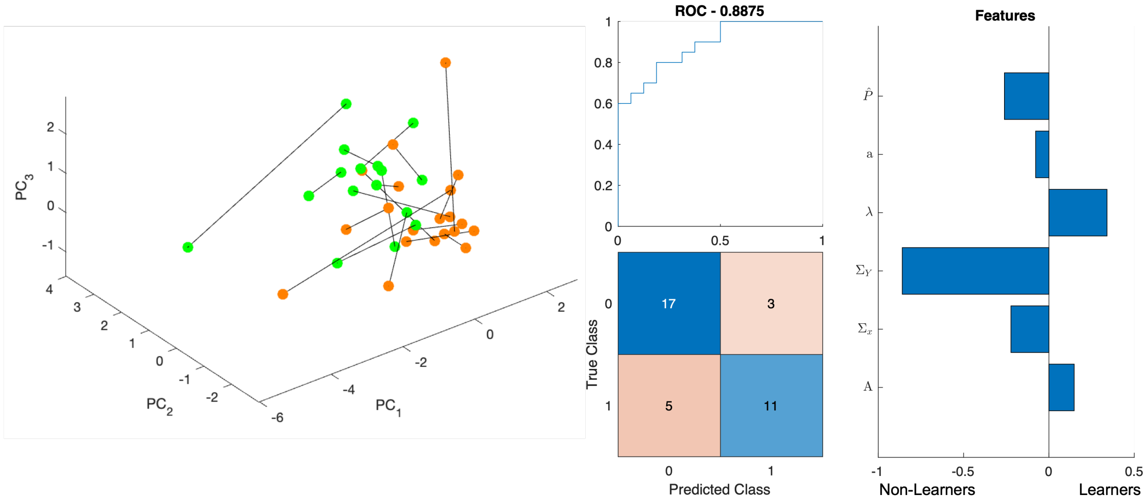

Figure 5.8 Determinants of learning. On the left, parameters for each players are projected in the principal component space. Each datapoint represents a player, each line connect players within the same dyad. In green are reported those players who are classified are Learner and in orange those who are classified as non Learner, as described in Section 2.3. Linear discriminant analysis is used to determine which are the parameters that characterize the different behaviors (Learner/Non Learner). Classification performances are reported in the confusion matrix and the ROC curve. On the right, Linear discriminant analysis (LDA) effect size (LEfSe) is reported. Horizontal bars represent the effect size for each feature - i.e. model parameter.

5.4 Discussion

We investigated how interacting agents negotiate a coordination over repetition in an interactive scenario in which achieving a coordination involves the selection of the appropriate strategy and spatiotemporal coordination within that strategy. To this purpose, we addressed a scenario in which agents share the same final goal but have different and incompatible intermediate sub-goals. This is particularly interesting since it encompasses all those situations in which agents work together to achieve the same goal but have different constraints that are not directly accessible and recognizable from the partner. This task is much more challenging than joint spatiotemporal coordination alone as studied by Chackochan and Sanguineti (2019), because to adeptly achieve a collaboration, players not only need to cross or approach partner's via-point, but they need to decide whether to do so before or after crossing own via-point. In other words, in this scenario even if both players within a dyad decide to collaborate, achieving a collaboration is more challenging, as they may never converge. To address this scenario, we manipulated the geometry of the 2-VP task: via-points are symmetrically located with respect to the horizontal axis, leading to two equally cost-effective strategies.

We observed a significant effect of time on spatiotemporal performance, which indicates that, at population level, the dyads increasingly coordinate their movements over trials. However, only a fraction of the dyads converged to stable coordination (ie. both players following either the 12 or the 21 strategies). Some dyads get stuck in cyclic behaviors, in which both players switch between 12 and 21 strategies. In other dyads, each player only focused on their own VP by ignoring the mechanical interaction with their partner, even though this implied more mechanical effort on their side. This latter finding suggests that even collaboration has a cost, which may overcome the extra mechanical effort required by lack of collaboration in this task.

We found no evidence that convergence to stable coordination depends on the reliability of information about the partner. This finding is in partial contrast with previous observations (Chackochan and Sanguineti, 2019), that more reliable sensory information has quantitative and qualitative effects on coordination, in the sense that the latter is faster and closer to the predictions of game theory (Nash equilibrium). However, although increasing the amount of information about the partner does not suffice to establish stable coordination, some dyads converged to a shared strategy. The question arises of whether some specific sensorimotor abilities and the peculiar aspects of partner representation facilitate this outcome.

A computational model of interaction – see Chapter 4 – is potentially useful to disentangle these observations. The joint task is represented as a quadratic game, and uncertainty is introduced at both perceptual and action selection levels. Each player predicts the actions of their partner and selects the action through a stochastic mechanism that accounts for the expected cost from that prediction. The player's sensory system provides information about the actions performed by both players, which is used to update their respective partner representations. We extended the model to account for the selection of the strategy by introducing a categorical variable s_i. We assumed that strategy selection is also dependent on partner representation. As a result, the action of each player has a Gaussian Mixture distribution in which strategy choice and the strategy-specific action distributions depend on the partner model x_i. The model is general enough to capture a variety of individual traits (e.g. high vs low reliability of sensory information; retention of partner representations).

The estimated model parameters did not capture the experimental manipulation of the sensory information (H and VH groups). This may be due to the fact the observed behaviors are a little different in the two experimental groups. It should be noted, however, that the relevant model parameter – sensory covariance Σ_y^i – reflects not only the experimental manipulation but also the individual sensory capability, which may vary in different individuals. However, when looking at Learners and Non-learners, we found that the former had lower sensory covariance Σ_y^i (i.e., better sensory capability) and greater initial temperature of action selection, λ^i (i.e., a higher degree of randomness in action selection, and therefore a greater tendency to exploration). These findings were confirmed by linear discriminant analysis, which also suggested that learners tend to have greater partner model retention A^i and smaller partner model randomness Σ_x^i. In other words, in learners, the partner model has more 'inertia' and relies more on previous history. Taken together, these features indicate that learners are less likely to modify their partner predictions and therefore to switch strategy from one trial to another. These observations suggest that we not only adjust our behavior to partner representations. But strategy choice may be subject to adaptation on its own. This recalls recent findings on motor learning (Hadjiosif et al., 2021) that suggested that adaptation may be driven not only by sensory-prediction errors, relying on internal forward models, but also on a learned policy that is adjusted over time by movement errors directly ("direct policy learning").

Chapter 6

Haptic communication in redundant sensorimotor tasks

6.1 Introduction

When addressing coordination between multiple individuals we can distinguish between planned and emergent coordination (Knoblich et al., 2011). In planned coordination, an individual's behavior is guided by representations about joint outcomes and own contributions to these outcomes. While emergent coordination relies on action-perception processes that are similar to multiple individuals. Emergent coordination (Knoblich et al., 2011) can occur between individuals who have no plans to act jointly, but also during planned joint actions. Emergent coordination may be understood as a facilitator of joint action (Vesper et al., 2010), as it seems to rely on a more low-level process (action-perception coupling) and does not require explicit reasoning about social interaction. To study emergent or unintended coordination one must create a situation in which the task only weakly constrains coordination (Schmidt and O'Brien, 1997). That is one in which participants' goal is other than the interpersonal coordination that results from interaction. Emergent coordination is usually addressed using rhythmic movements and different perceptual conditions. In the study of Schmidt and O'Brien (1997) each participant within a dyad controlled the motion of a hand-held pendulum. Their goal was to swing the pendulum at a comfortable tempo and maintain it while viewing the other person's rhythmic movements. They found that participants did not become frequency-locked when viewing the other person's pendulum, indicating that participants followed task instructions. However, they found that the relative timing of the movements was affected by the visual coupling and the correlation of the two movements increased indicating that entrainment occurred. Further, they interestingly modeled unintended interpersonal rhythmic coordination with oscillatory equations that have been used to model intended interpersonal rhythmic coordination (Haken et al., 1985). van der Wel et al. (2011) addressed haptic-mediated coordination in a task where interacting agents shared the control of a pendulum motion. They found that dyads amplify their forces to generate haptic information channel. Oullier et al. (2008) found that spontaneous phase synchrony (i.e., unintentional in-phase coordinated behavior) between two people emerges as soon as they exchange visual information, even if they are not explicitly instructed to coordinate with each other.

In neuroscience, motor learning has most often been studied in the context of adaptation of reaching movements (Stockinger et al., 2014). Thereby, subjects usually adapt their reaching movements to either kinematic perturbations (visuomotor rotations, (Krakauer et al., 2005) prism-induced displacements (Held and Freedman, 1963)) or dynamic perturbations (robot-induced forces (Shadmehr and Mussa-Ivaldi, 1994); rotation of body, (Lackner and DiZio, 2005); attached inertial loads (Krakauer et al., 1999)). Here, we focus on motor learning in terms of adaptation of reaching movements to human-induced forces, instead of robot-induced forces. Thereby, subjects interact with a robotic device that applies perturbing forces to the subjects' hand leading to changed dynamic conditions of the reaching movements. The applied forces do not depend on the fixed pre-determined task-dependant physical laws but depend on the real-time movements of the partner.

Point-to-point reaching task has been addressed also in physical human-human interaction (Reed and Peshkin, 2008; Takagi et al., 2016). Reed and Peshkin (2008) argued that interacting agents exploit haptic communication to specialize their contribution into the shared task. Instead Takagi et al. (2016) argued that interacting agents did not coordinate their motion. Variant of reaching movement has been proposed to address collaboration in tasks where interacting agents mechanically coupled had partly incompatible goals (Chackochan and Sanguineti, 2019). Collaboration is achieved as information reliability about the partner increases (Chackochan and Sanguineti, 2019). Coordination in goal-directed task is a matter of debate. To this end tracking task (Ganesh et al., 2014; Melendez-Calderon et al., 2015; Takagi et al., 2017, 2018) has been proposed as more suitable to explore coordination, as allows for longer movements, however it constrains participants to the target timing and leaves little freedom of motion in term of trajectory selection. In the 2-via-point task proposed by Chackochan and Sanguineti (2019) and addressed in the previous chapters, timing is constrained by the feedback provided at the end of each trial.

While emergent coordination is a general focus of the current thesis, here we specifically intended to explore more ecological scenarios. We address a setting that involves a variety of possible coordination patterns and no time constraints. To this purpose, we devised redundant sensorimotor tasks, where participants are required to perform complex reaching movements (Kodl et al., 2011; Todorov and Jordan, 1998) while mechanically coupled with a virtual spring. Redundancy is given by the freedom of players to choose their preferred solution in terms of both space and time. We present preliminary results from two sets of experiments. We tested pairs of participants who were required to perform reaching movements using an end-effector robot. Each participant controlled the motion of a cursor on the screen in from of him/her. Robots were programmed to render a virtual spring connecting the two handles.

In the first experiment, participants were instructed to perform reaching movements through the same sequence of via-points and received the same visual cue in the initial phase of the experiment. The visual cue displayed a possible trajectory to be used to perform the task. After this initial phase players were mechanically coupled and received no prior visual cue or feedback at the end of the motions. After a long training while connected through a virtual spring, players returned to perform the task in solo condition for a few after-effect trials.

In the second experiment, we manipulated the initial visual cue. The task was the same for both players within a pair - they had to reach the same sequences of via-points - but they were initially instructed with two different visual cues. With this second experiment, we address a scenario where players have different prior knowledge on how to perform the same task and have suddenly to perform the same task while mechanically coupled. We assess the level of coordination in terms of temporal synchronization and trajectory similarity metrics. Overall, we found that participants changed their behavior when mechanically coupled with respect to their solo performances. When mechanically coupled, participants changed their behavior (toward coordination). However, coordination rapidly disappears when the mechanical coupling is removed. These results provide insights into the understanding of the mechanisms underlying the development of a partner model and its contribution to the action selection process.

6.2 Materials and Methods

6.2.1 Experimental apparatus and task

Each experiment involved one pair of participants (a dyad). Participants sat in front of two separate computer screens and grasped the handle of a planar haptic interface (H-MAN Campolo et al. (2014)). Each participant's seat position and height were adjusted so that they were able to comfortably reach any position of the workspace of the robot. They could not see or hear each other, and they were not allowed to talk. The experimental apparatus is depicted in figure 6.1a. The setup is described in detail in section 6.4. Participants were instructed to perform reaching movements in the horizontal plane. By moving the handle of the robot they controlled the motion of a cursor on the screen. Participants were instructed to stay still on a start position until a ’Go!’ signal appeared and to perform movements as continuous and smooth as possible in a way that was comfortable to them from the start to the end crossing the via-points. No feedback on performance was given to the subject at any stage of the experiment.

6.2.2 Experiment 1

Subjects 10 naive subjects (7 females) without known pathology, aged between 20 and 31, participated in the experiment and were randomly grouped into pairs. The experiments were conducted according to the principles in the Declaration of Helsinki and were approved by the ethics committee at Imperial College London. Each subject gave informed consent prior to involvement in the study.

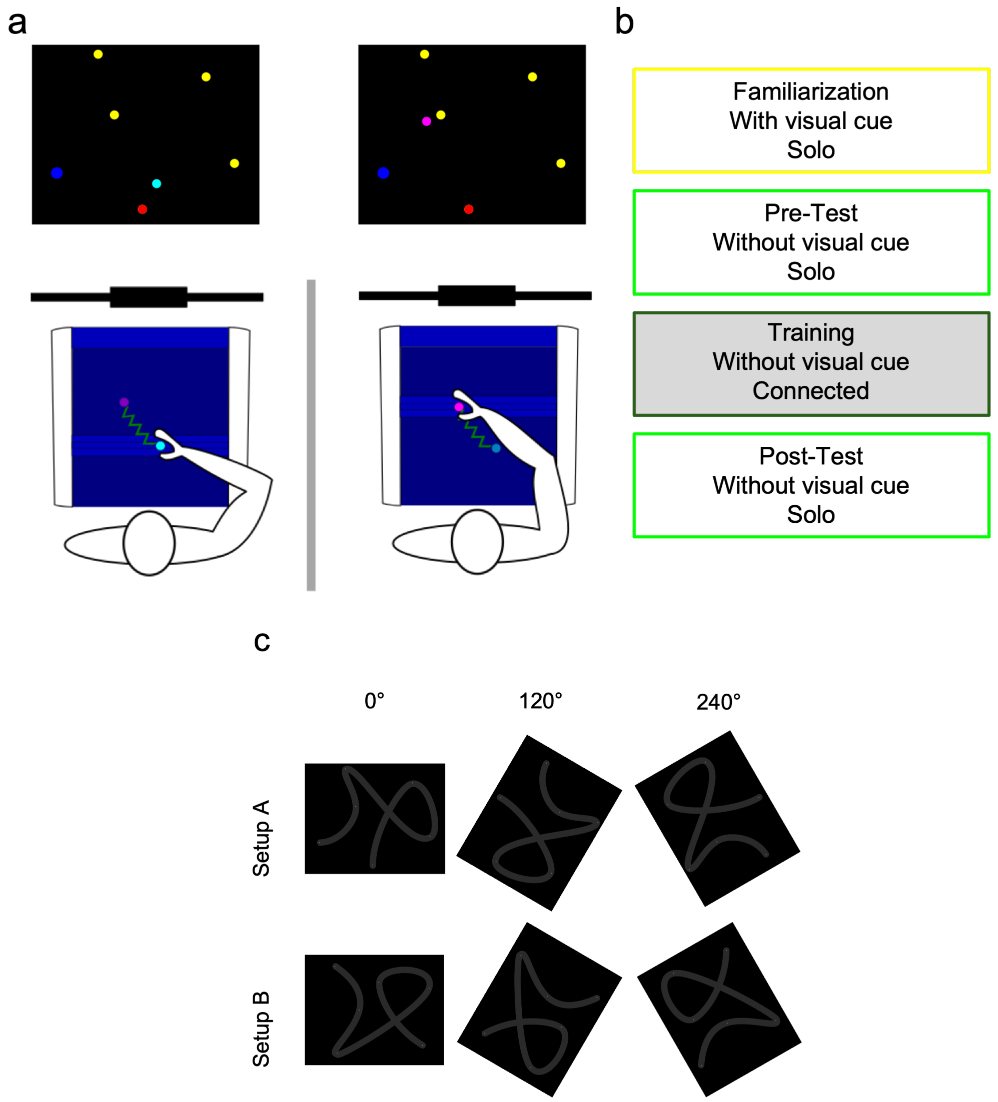

Figure 6.1 Experiment 1. a) Two participants sat in front of a computer screen and a planar manipuladum. Interacting participants were separated by a curtain and could not see or hear the each other. They were required to perform reaching movement using the haptic device from a start point (blue), by crossing a sequence of numbered via-points (yellow), and finish their movement at the end point (red). They were required to perform natural and comfortable movements as smooth as possible with no timing constraints. b) The protocol consists of a familiarization phase during which players received the same visual cue before the 'Go!' signal about a possible trajectory they could use to carry out the task. In the pre-test, players perform the same task with no visual cue. The familiarization and pre-training test lasted 16 trials and players perform the task with no mechanical coupling. During the training phase (84 trials) players had to perform the same task, with no visual cue while mechanically coupled. During the last test phase (12 trials) players were not mechanically coupled and received no initial visual cue. c) Trajectories were varied randomly during the experiment. In particular, 2 via-point configurations were rotated at 0°, 120°, and 240°.

Experimental protocol In this first experiment, the location of the start, via-points, and end point varies along the experiment trial by trial but was the same for the two participants. In particular, we used two different setups and three different rotations as in Kodl et al. (2011). No feedback on performance was given to the subject at any stage of the experiment. The protocol is reported in Figure 6.1b. In a first familiarization phase (4 trials) participants received a visual cue representing a possible trajectory they could use to solve the task. The displayed trajectory was the same for the two players. For the next 12 trials (Pre-test phase) they no longer receive the visual cue about the possible solution of the task. In these two initial phases, players were not mechanically coupled. This means that they were performing their motions with no constraints from the partner. After this phase the mechanical coupling was activated and they had to perform a training phase of 84 trials. Within the training, catch trials were included with a ratio 1:6 in which the mechanical coupling was deactivated. After training players had to perform a Post-test including 12 trials, in which they had to perform the task with no force. The experiment lasted approximately 45 minutes. Participants were initially required to fill the consent form and answer a demographic questionnaire.

6.2.3 Experiment 2

Subjects 14 naive subjects (7 females) without known pathology, aged between 21 and 35, participated in the experiment and were randomly grouped into pairs. The experiments were conducted according to the principles in the Declaration of Helsinki and were approved by the ethics committee at Imperial College London. Each subject gave informed consent prior to involvement in the study.

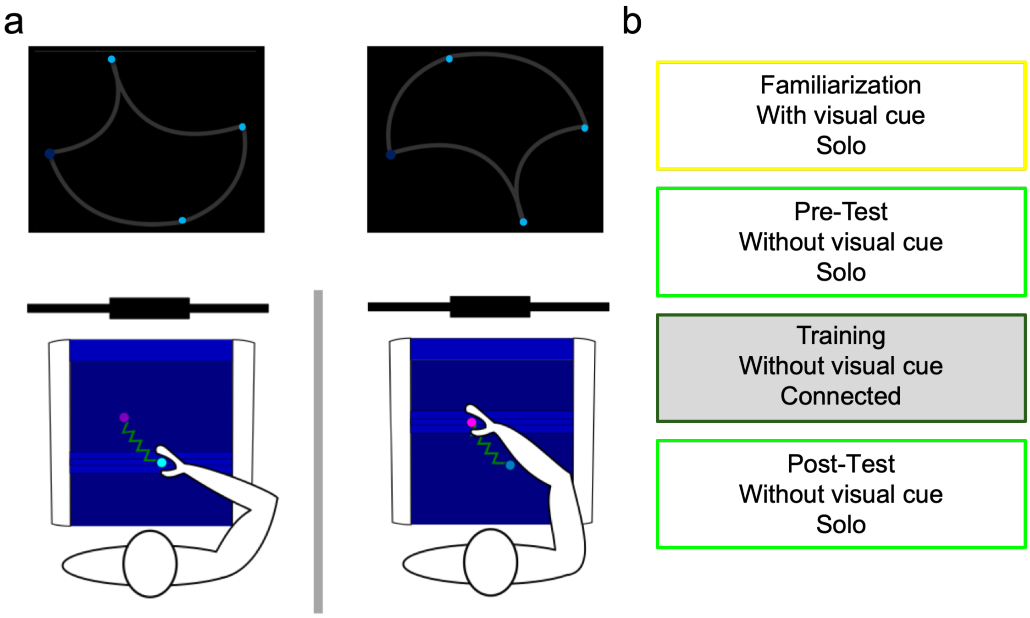

Figure 6.2 Experiment 2. a) Two participants sat in front of a computer screen and a planar manipuladum. Interacting participants were separated by a curtain and could not see or hear the each other. They were required to perform reaching movement using the haptic device from a start point (blue), by crossing a sequence of via-points (turquoise) and returning to the start. They were required to perform natural and comfortable movements as smooth as possible with no timing constraints. b) The protocol consists of a familiarization phase during which players received a visual cue before the 'Go!' signal about a possible trajectory they could use to carry out the task. In the pre-test, players perform the same task with no visual cue. Familiarization and pre-test lasted 15 trials and players perform the task with no mechanical coupling. During the training phase (120 trials) players had to perform the same task, with no visual cue while mechanically coupled. During the last test phase (12 trials) players were not mechanically coupled and did not receive the initial visual cue.

Experimental protocol In the second experiment the location of the start, via-points, and end point was fixed, and participants were required to perform clockwise movements. The protocol is summarized in Figure 6.2b. In the first familiarization phase (12 trials) participants could see a possible trajectory they could use to solve the task. Each participant received a different visual cue: the suggested trajectories differed in the curvature of the trajectory sub-segments. During movements, they could only see the via-points on the workspace. For the next 4 trials (Pre-test phase) they no longer receive the visual cue about the possible solution of the task. In these two initial phases there was no connection. This means that they were performing their motions with no constraints from their partner. After this phase, the mechanical coupling was activated and they had to undergo a 120-trial training phase. Within the training, catch trials were included with a ratio 1:6 in which the mechanical coupling was deactivated. After training players had to perform a post-training test including 12 trials, in which they had to perform the task with no force. The experiment lasted approximately one hour. Participants were initially required to fill out the consent form and answer a demographic questionnaire.

6.2.4 Data Analysis

During experiments, hand trajectories were sampled at 100 Hz and stored for subsequent analysis. Movement trajectories were smoothed by means of a 4th order Savitzky-Golay filter with a 370 ms time window. The same filter was used to estimate velocity and acceleration. For each individual trial and for each player within a dyad, we identified the start and end times of each trajectory, as the time instants ST_i and ET_i at which the speed of player i crossed a 1 cm/s threshold.

Coordination can be characterized in terms of movements kinematics and kinetics, at both individual and dyad level. Full coordination implies alignment in space and time, but in principle, interacting agents may coordinate their action in space but not in time, or in time but not in space. We are interested in understanding whether coordination in space and coordination in time result from the same process or from two parallel processes running at different rates and leading to a different proportion of coordination. To this purpose, we decided to use metrics that allow disentangling spatial and temporal coordination - i.e. timing differences and shape similarities.

A requirement for coordination is that players within a dyad start their movement together. We observed the difference of start time of the two players within a dyad, $\Delta\mathrm{ST} = |\mathrm{ST}_1 - \mathrm{ST}_2|$. Coordination requires not only synchronization of the start time but also the times of crossing via-points should be aligned. To this purpose, for each player i we identified the via-point crossing time, TC_{ij} as the time instant in which the distance from via-point j is minimal. We quantified the Minimum Via-Point Distance, MD_{ij} of player i from via-point j: $\mathrm{MD}_{ij} = \min_t \mathrm{dist}_{ij}(t)$ with $\mathrm{dist}_{ij}(t) = ||\mathrm{x}_i(t) - \mathrm{x}_{\mathrm{VP}_j}||$, $i = 1,2$ and $j = 1,...,\mathrm{N}_{VP}$. Where N_{VP} depends on the experiment - N_{VP} = 4 for Experiment 1 and $\mathrm{N}_{VP} = 3$ for Experiment 2. We observed the difference of crossing time at each via-point j, $\Delta\mathrm{TC}_j = \left|\mathrm{TC}_{1j} - \mathrm{TC}_{2j}\right|$. Coordination implies that these differences are close to zero. Conversely, if interacting participants ignore each other and use their own timing the difference is higher. As a synthetic measure of temporal synchronization, we considered the average of the differences of crossing time at each via-point. Coordination is not just a matter of temporal synchronization. To adeptly coordinate with others we need to adopt similar trajectories. To evaluate trajectory similarity we considered Dynamic Time Warping (DTW) (Vintsyuk, 1968). DTW measures the distance between two temporal sequences. This algorithm stretches two vectors onto a common set of instants such that the distance, the sum of the Euclidean distances between corresponding points, is the smallest. Low values of this metric indicate that trajectories are similar in space.

Another metric addressing coordination in space and time is the speed profiles correlation. We considered the movements of each player from the first start time - $\mathrm{ST} = \min(\mathrm{ST}_1, \mathrm{ST}_2)$ - to the last end time - $\mathrm{ET} = \max(\mathrm{ET}_1, \mathrm{ET}_2)$ - and computed the correlation coefficient between the speed profiles of the two players. Values close to 0 indicate that interacting participants are not accounting for each other, while values close to 1 indicate that partners' movements are strongly aligned.

Haptic-mediated coordination implies that interacting participants understand the partner and the dyad dynamic through forces exchange and attempt to reduce the effort, trying to minimize the interaction force. To this purpose we observed the interaction force for each trial and we computed the average interaction force $IF = \frac{1}{N} \sum_t \|F(t)\|$, where $F(t)$ is equal and opposite for the two partners in the dyad.

Statistical Analysis We expect that task performances at players and dyad level evolve with time. To this purpose, we tested whether training while mechanically coupled would lead to a change in uncoupled performance, by comparing pre-training and post-training unconnected trials. We averaged the last six trials in the unconnected pre and post-training tests for each dyad and performed a one-way ANOVA. If data were not normally distributed, we used the non-parametric Kruskall-Wallis test. To address the change in dyad performance during training we considered, for each dyad, the average of the last three connected trials and the unconnected trial for each epoch. In particular, to address performance evolution and the effect of the mechanical coupling during training we perform a two-way repeated measure ANOVA, with Time and Connection as within factors. If data were not normally distributed we used the non-parametric Friedman test with Time and Connection effect.

Data distribution was first examined using Kolmogorov-Smirnov test. The significance threshold was set at 0.05.

6.3 Results

With these experiments, we intended to explore if and how haptic coupling affects performances in a multi-via-points reaching task. In particular, we were interested in understanding whether training while mechanically coupled lead to a change in performances in the connected trials and in the unconnected trials.

6.3.1 Experiment 1 - Collaboration

In the first experiment, we aim at testing whether physically coupled participants coordinate their movements when performing a multi-target reaching task. And if they coordinate, to which extent they do so when they share the same goal. Overall, we observed that, when participants are mechanically coupled through a virtual spring they tend to synchronize their motions in space and time. A summary of the statistical results is reported in Table 6.1.

	Training phase			**Unconnected phases**
Metric	Time	Connection	Time × Connection	Pre-test vs Post-test
Asynchrony	NS	0.012	NS	0.0006
DTW	NS	0.0005	NS	0.0005
Speed Correlation	NS	0.0002	NS	0.0037
Interaction Force	NS	-	-	-

Table 6.1 Statistical analysis for Experiment 1. For each metric p-value are reported. NS stands for Non Significant.

This can be observed qualitatively from the trajectories and the speed profiles in figure 6.3.

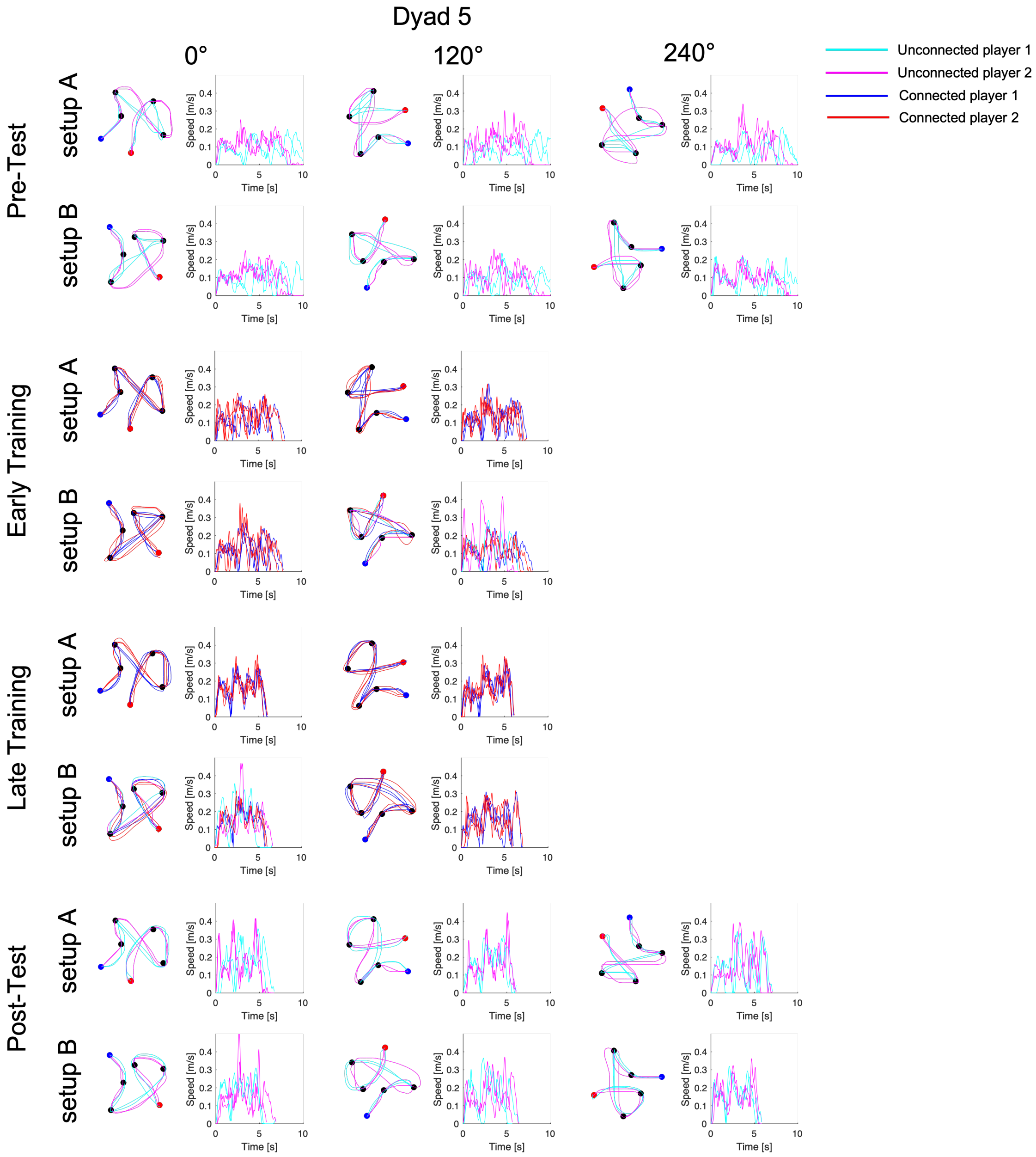

Figure 6.3 Trajectories and speed profiles from one representative dyad. In magenta and turquoise unconnected trials are reported, in red and blue trials with the active mechanical. Trajectories are reported for the two via-points configuration and for the three (in pre- and post-tests) and two (during training) rotations. Blue and turquoise are used to represent trajectories and speed profiles of player 1 and red and magenta are used to represent player 2.

We observed asynchrony to address temporal synchronization, as it is one major aspect of coordination.

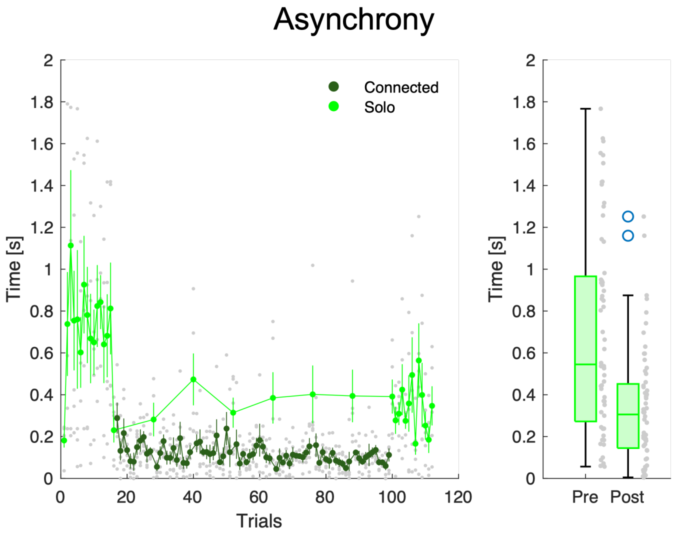

Figure 6.4 Asynchrony is computed as the average of asynchrony at the start and the different VPs. Left) For each trial performances for each dyad are reported in grey. In green the average and SE at population level are reported. Trials in which mechanical coupling was active are reported in dark green. Trials in which mechanical coupling was removed are reported in light green. Right) we reported box plots for epochs 2 and 10, corresponding to the pre and post-test phases.

Asynchrony decreases after training in the unconnected trials (p = 0.0006, F(1,58) = 13.0950). Further, we found no effect of Time or interaction Time × Connection, but we found a Connection effect (p = 0.012, F(1,58) = 57.8). These results indicate that players at the beginning of the experiment (pre-test) used a timing that is different from the partner's timing. During the training phase, connection strongly affects synchronization. Performances do not seem to change with time during the training phase, however, at the end of the experiment, players are more temporally aligned.

When looking at coordination in space we found that Dynamic Time Warping significantly changes with time when looking at the unconnected pre and post-tests (p = 0.0005, F(1,58) = 13.5302). As before, we found that Connection significantly affects performances during the training phase (p= 0.0005, F(1,16) = 18.76), but we found no effect of Time or interaction Time × Connection. This indicates, similarly to what we found for the asynchrony, that players select a strategy for the connected trials which is stably different from the one selected in the unconnected catch trials.

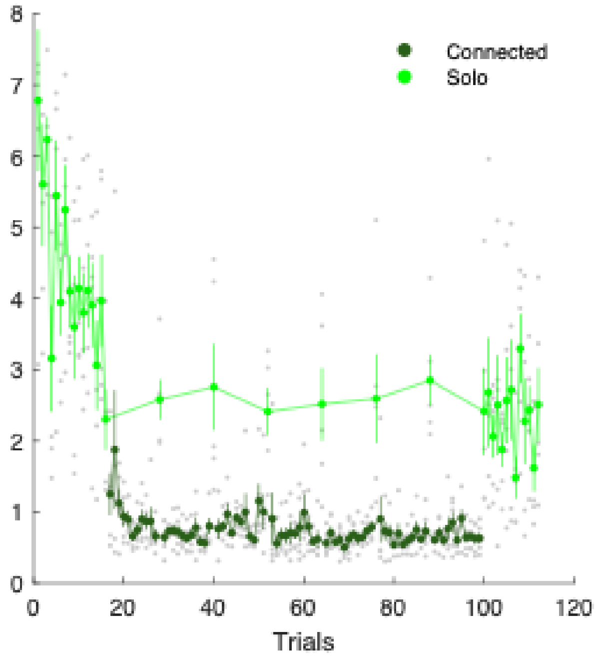

Figure 6.5 Dynamic Time Warping. Left) For each trial performances for each dyad are reported in grey. In green the average and SE at population level are reported. Trials in which mechanical coupling was active are reported in dark green. Trials in which mechanical coupling was removed are reported in light green. Right) we reported box plots for epochs 2 and 10, corresponding to the pre and post-test phases.

When looking at speed correlation, we found that time affects this metric when comparing pre and post-tests ($p = 0.0037$, $F(1,58) = 9.1252$). We also found a strong Connection effect on performances during the training phase ($p = 0.0002$, $\chi^2(1,16) = 13.64$) but no Time effect.

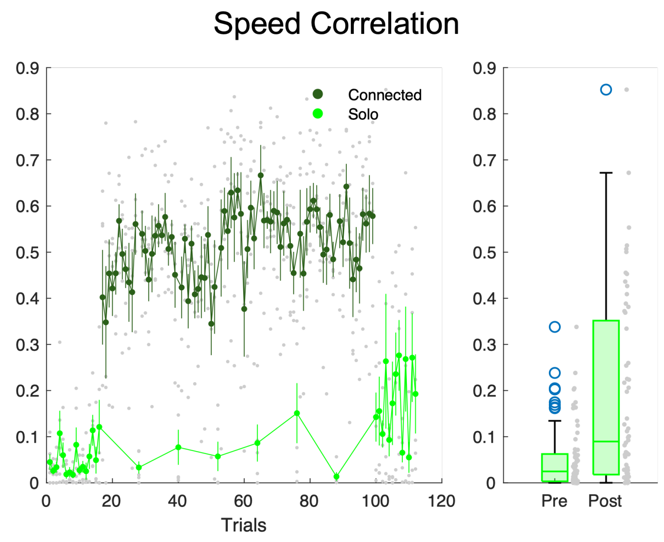

Figure 6.6 Speed correlation. Left) For each trial performances for each dyad are reported in grey. In green the average and SE at population level are reported. Trials in which mechanical coupling was active are reported in dark green. Trials in which mechanical coupling was removed are reported in light green. Right) we reported box plots for epochs 2 and 10, corresponding to the pre and post-test phases.

As regards effort, we found that interaction force between players does not decrease with training.

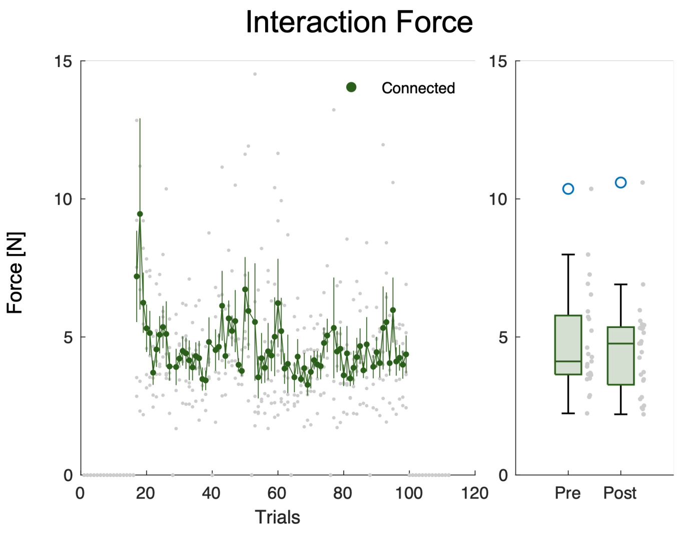

Figure 6.7 Interaction Force. Left) data are reported for each dyad in grey. The average value is reported for each trial in dark green. Right) box plots are reported for epochs 3 and 9, corresponding to the first and last connected epochs.

6.3.2 Experiment 2 - Negotiation

In the second experiment, we aim at testing a similar scenario, but we were interested in understanding whether physically coupled players, with the same task, coordinate their movements when they receive different initial visual cues. This setting address all these scenarios where two agents individually know how to solve the task and their solutions do not match. The sequence of via-points was the same for both players but the cued paths were different which implies that different possible time courses. Participants were instructed to perform arm movements crossing a sequence of via points. Participants within a dyad had identical via-point sequences but received different initial visual cues. We expected that players converged toward a shared trajectory, minimizing their effort due to the interaction, as the only requirement was to cross a sequence of via-points which was identical for the two players. A summary of the statistical results is reported in Table 6.2.Players while mechanically coupled tended to use a similar timing, but they do not use the same trajectory. Instead, they tend to use their own initial solution. When the connection was suddenly removed, they instantaneously return to use their individual trajectory. Players did not completely minimize their effort, instead, they preferred to use the trajectory they were shown at the beginning of the experiment even though they were not required to. This can be observed qualitatively from the trajectories and the speed profiles in Figure 6.8.

	Training phase			**Unconnected phases**
Metric	Time	Connection	Time × Connection	Pre-test vs Post-test
Asynchrony	NS	NS	NS	$p < 0.0001$
DTW	NS	0.0299	NS	0.0006
Speed Correlation	NS	0.0001	NS	NS
Interaction Force	0.0309	-	-	-

Table 6.2 Statistical analysis for Experiment 2. For each metric p-values are reported. NS stands for Non Significant.

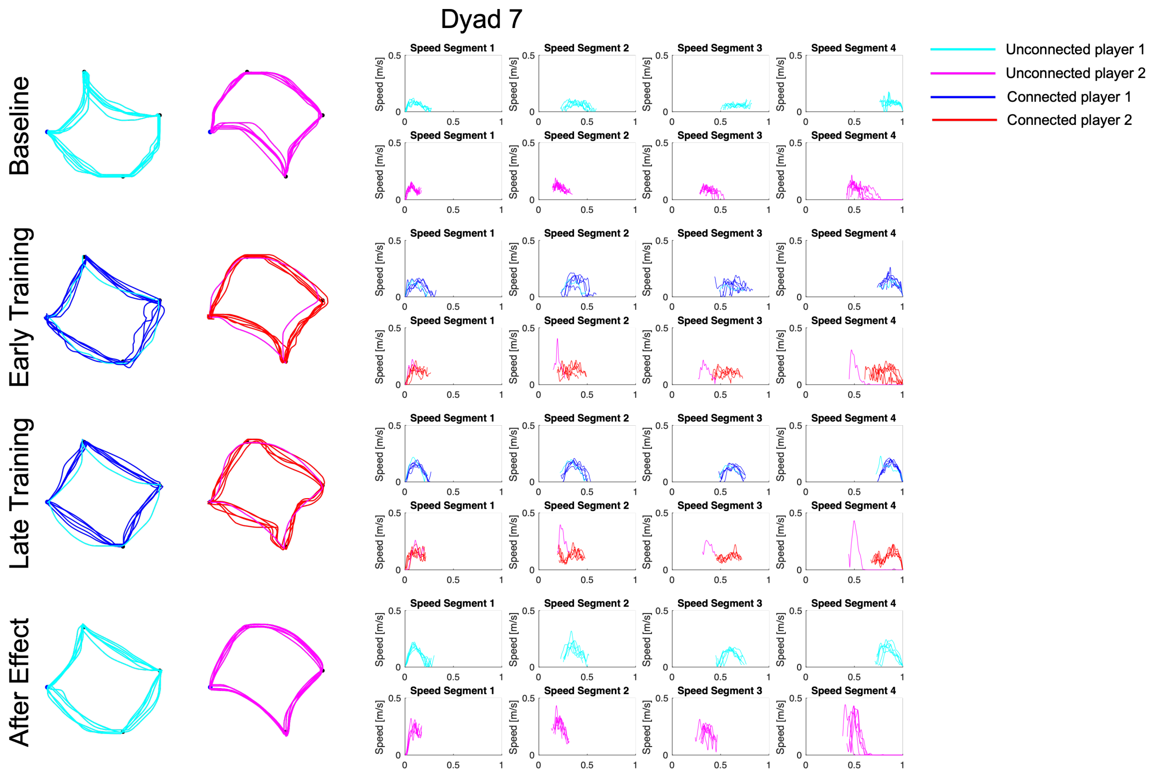

Figure 6.8 Trajectories and speed profiles for one representative dyad. In magenta and turquoise unconnected trials are reported, in red and blue trials with the active mechanical coupling. Speed profiles are divided by movement segment. Blue and turquoise are used to represent the trajectories and speed profiles of player 1 and red and magenta are used to represent player 2.

As for the previous task, temporal synchronization is crucial for coordination. Start time asynchrony changes considering pre- and post-tests ($p < 0.0001$, F(1,82) = 17.3706). Further, we did not find a significant Time or Connection effect or interaction Time × Connection in training performances.

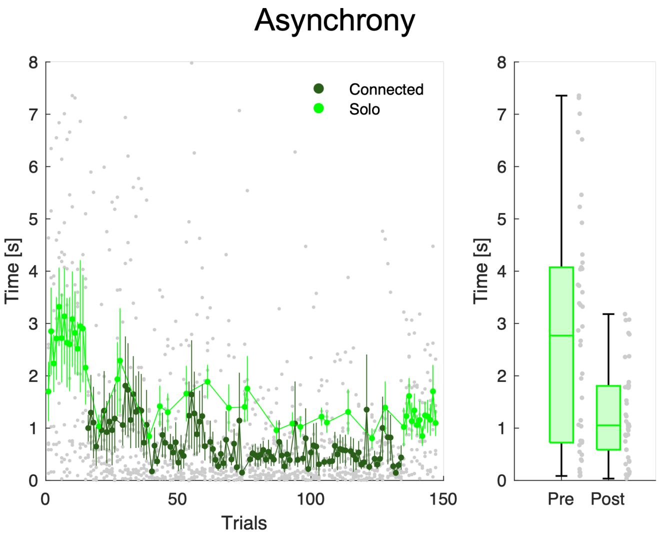

Figure 6.9 Asynchrony is computed as the average of asynchrony at the start and the different VPs. Left) For each trial performances for each dyad are reported in grey. In green the average and SE at population level are reported. Trials in which mechanical coupling was active are reported in dark green. Trials in which mechanical coupling was removed are reported in light green. Right) we reported box plots for epochs 2 and 25, corresponding to the pre and post-test phases.

Addressing trajectory similarity, we found that DTW we found a significant Time effect comparing pre and post-tests (p = 0.0006, F(1,84) = 12.54). Looking at performances during the training phase, we found no significant effect of Time or interaction Time × Connection but a significant effect of Connection (p = 0.0299, F(1,24) = 5.33).

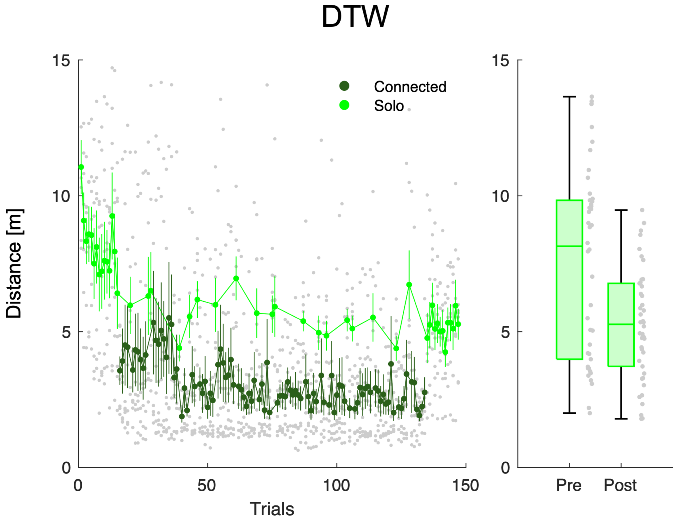

Figure 6.10 Dynamic Time Warping. Left) For each trial performances for each dyad are reported in grey. In green the average and SE at population level are reported. Trials in which mechanical coupling was active are reported in dark green. Trials in which mechanical coupling was removed are reported in light green. Right) we reported box plots for epochs 2 and 25, corresponding to the pre and post-test phases.

When looking at speed correlation we found no significant difference when comparing pre and post-tests. We found no significant effect of Time or interaction Time × Connection, but we found an effect of Connection ($p = 0.0001$, $F(1,24) = 22.58$).

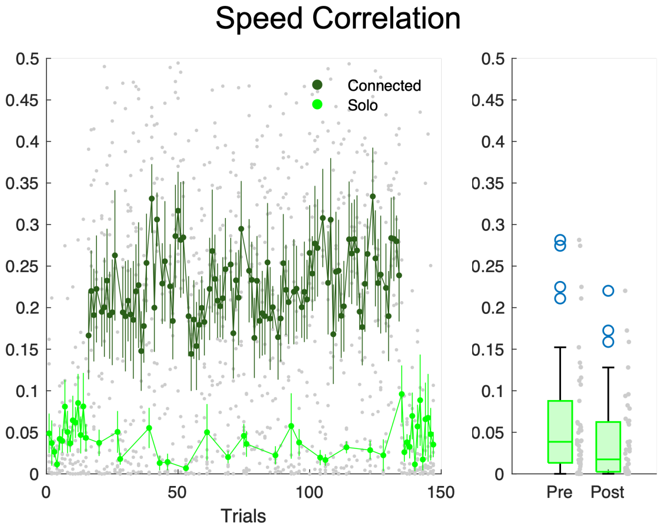

Figure 6.11 Speed correlation. Left) For each trial performances for each dyad are reported in grey. In green the average and SE at population level are reported. Trials in which mechanical coupling was active are reported in dark green. Trials in which mechanical coupling was removed are reported in light green. Right) we reported box plots for epochs 2 and 25, corresponding to the pre and post-test phases.

As regards effort, we observed that interaction force significantly decreases with training ($p = 0.0309$, $F(1,68) = 4.8613$).

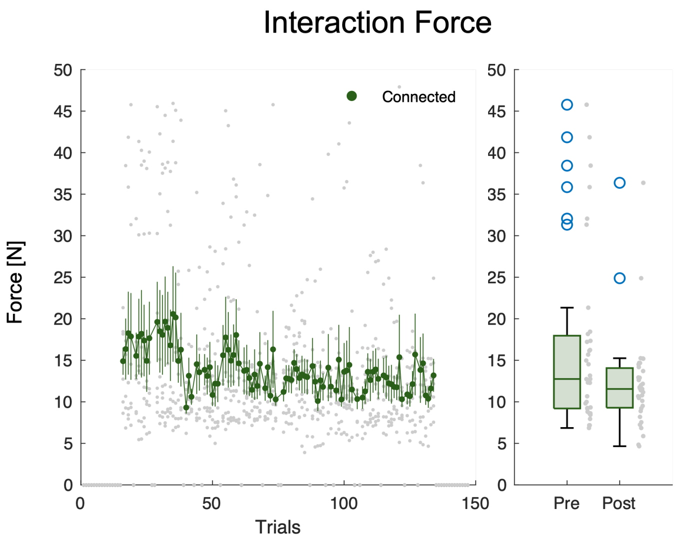

Figure 6.12 Interaction force. Left) For each trial performances for each dyad are reported in grey. In green the average and SE at population level are reported. Trials in which mechanical coupling was active are reported in dark green. Trials in which mechanical coupling was removed are reported in light green. Right) we reported box plots for epochs 4 and 23, corresponding to the pre and post-test phases.

Overall, these results indicate that coordination is affected by mechanical coupling. Players tend to use more similar trajectories when they are mechanically coupled. This is interesting as it seems that players account in some measure for the partner action, however, they do not learn as in the individual learning paradigm (force-field adaptation experiments). Here players instantaneously co-adapt to the partner as the coupling is activated, but equally instantaneously they tend to switch to their initial strategy.

6.4 Discussion

The study aimed to investigate coordination in physical human-human interaction in ecological and natural arm movements. In particular, we focused on a scenario where mechanically coupled agents have the same or different prior knowledge of possible strategies to perform the task. We were interested in understanding whether interacting agents used the same strategy or not and whether they changed their behavior while training together and retain the learned strategy in the subsequent solo phase. To this purpose, we used a dyadic experimental setup consisting of a pair of planar hand manipulandum which was programmed to render a virtual spring connecting the two interacting agents. With respect to previously proposed coordination tasks, we put no temporal (Chackochan and Sanguineti, 2019; Takagi et al.,

2016) or spatiotemporal constraints (Ganesh et al., 2014; Melendez-Calderon et al., 2015; Takagi et al., 2017) in action performances.

In the first experiment, we used a multi-via-points reaching task where individuals are required to perform reaching movements by crossing a number of via-points in a predetermined order. Participants were initially shown a possible solution, they were tested in solo condition, then they were mechanically coupled through a virtual spring for a long training phase and then tested again in solo condition. Participants within a pair had the same task and received the same visual cue at the beginning of the experiment. We found that players coordinate their movements as the connection is applied, their behavior did not evolve with time in the training phase, while mechanically coupled, nor they reduced their interaction force. However, in the training phase, players instantaneously changed their behavior in the unconnected catch trials. Further, after training the performances changed with respect to the initial performances in the same unconnected condition. To test whether interacting partners dynamically negotiate a shared strategy through training while haptically coupled we tested a similar scenario where agents have the same goal but have different prior knowledge about how to solve the task. In this experiment, players received different visual cues of possible trajectories they can use to carry out the task. This resembles a situation where people are physically connected and have a possibly different preferred solution on how to solve a task and need to negotiate a shared strategy. Interestingly, we found qualitatively similar results to the first experiment. Players' behavior is instantaneously affected by mechanical coupling and there is an effect of time when comparing performances in solo trials before and after training. We expected a gradual convergence toward a shared strategy. In this latter experiment, we found a slight decrease in interaction force. However, the evolution during training is not significant for the other coordination metrics.

These experiments indicate that physically interacting agents coordinate their motions and reduce their effort to some extent but did not incorporate an internal model of the perturbation as in force field adaptation or visuorotation learning experiments. In those experiments, initial errors were gradually corrected by incorporating an internal representation of the disturbance in the motor plan. This was observed when the disturbance was suddenly removed and participants had an error opposite to the initial one under disturbed conditions. Here instead, we found that during unconnected trials of training the players changed their behavior toward their individual preferred solution. This suggests that while interacting with others we build an internal representation of the partner which is instantaneously incorporated into the action selection process but is weakly retained when performing the same task in isolation. More in general the presented experiments may indicate that the mechanisms underlying learning of coordination are different from the mechanisms underpinning individual sensorimotor learning.

These experiments point out that haptic-mediated coordination is not complete and that players stably accept a degree of effort due to the interaction force. In terms of optimal control, these results may indicate that there is a trade-off between error and effort. Effort related to the interaction is counterbalanced by the accuracy of their initial preferred strategy. The observation that players tend to return to their preferred strategy during the catch trials indicates that they do not change their motor plan with training and suggests that players interpret interaction force as a disturbance. These observations are consistent with other experiments (Chackochan and Sanguineti, 2019) where it has been found that people do not completely achieve a shared strategy if coupled only through the haptic channel. Indeed in Chackochan and Sanguineti (2019) players converged toward the optimal joint solution - i.e. the Nash equilibrium - only when visual information about the partner was added to the haptic coupling. Altogether, these results may indicate that haptic coupling may not suffice to establish full joint coordination.

Interestingly, these results may be explained in terms of the computational model proposed in Chapter 2 to describe the sense of agency in joint action. The model posited two sensorimotor control loops which run in parallel continuously evaluating the assumptions of acting alone or with someone else. The model allowed us to predict that as information about a possible partner becomes available, the experience of agency shifts toward a sense of joint agency, accordingly determining a change in the action to be issued. In the experiments presented in this Chapter, we found that strategies in the catch trials instantaneously change toward individual preferences. Overall, these findings suggest that information about the partner and internal representation of the partner affect our behavior. In particular, in previous Chapters 4 and 5 we focused on learning through repetition in a scenario where players have partly different goals. Here, where players have the same goal but different preferences, it seems to emerge a different dynamic of coordination. Overall this work highlights the importance of visual cues in physical human-human interaction and provides insights into delineating the dynamics which govern the emergence of coordination.

Experimental set-up for human-human interaction

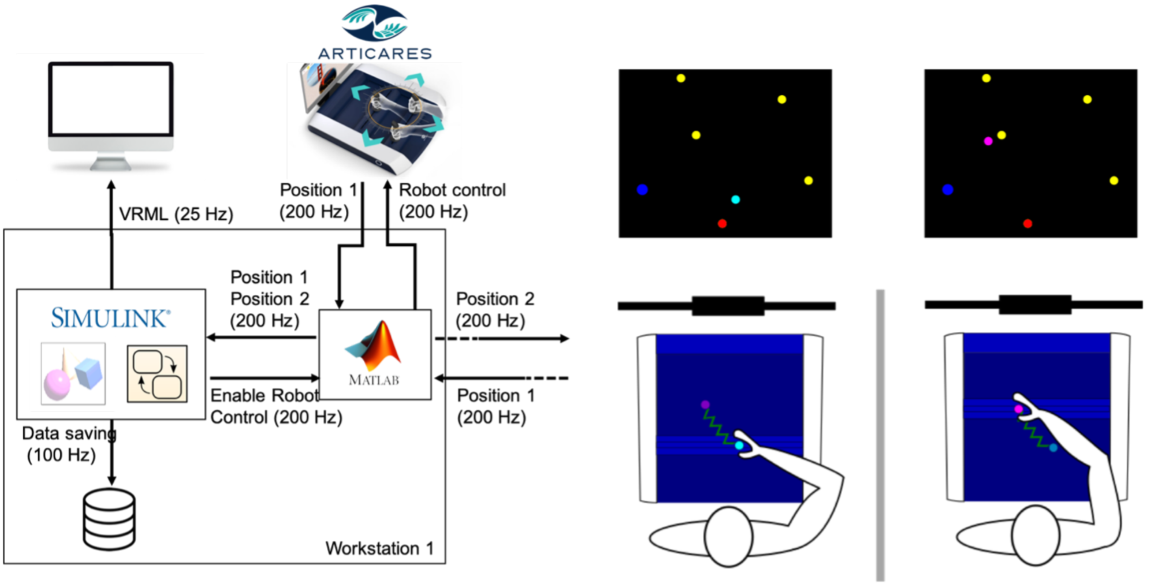

Figure 6.13 Left) Software architecture: this represents one side of the dyadic architecture. Each robot is connected to a computer (workstation). On the workstation a Matlab application and a Simulink model run in parallel. Right: Dyadic setup: each participant seat in front of a computer screen and grasp the handle of a robotic interface. Participants are mechanically coupled through a virtual spring and instructed to perform planar arm movements.

This section describes the experimental setup developed for the experiments described in the current chapter.

We developed an experimental set-up based on two identical 2-DoF haptic interfaces (H-Man), each with a dedicated control computer and a 27-inch LCD monitor for the presentation of visual feedback. The H-Man (Campolo et al., 2014) is a planar manipulandum, manufactured by Articares for clinical settings and home rehabilitation (Click This). The main advantage of this system is its lightweight mechanics, characterized by simplified kinematics and intrinsic high isotropy over the workspace. H-Man is characterized by high manipulability, dexterity, and isotropy over the entire workspace. This results from simple kinematic topology where two actuators in differential configuration drive two perpendicular linear sliders upon which the end-effector is located. The force transmission from the actuators to the end-effector is based on an H-shaped cable-driven transmission system. A prototypical version of the planar manipulandum has been used in previous works addressing robotic assistance thresholds in stroke rehabilitation (Kager et al., 2019a) and in the joint scenario, the effect of skill matching between interacting partners on learning a visuomotor task (Kager et al., 2019b). H-Man is connected through the control computer through an ethernet cable using TCP/IP communication protocol. The commands sent by the computer through the interface are interpreted by the integrated firmware. Similarly, the data recorded by the optical encoders are transmitted back to the control PC. This interface uses a nominal refresh rate of 200 Hz. To develop the dyadic set-up, allowing the exchange of data between the two H-Mans we used two different control computers, one for each device that exchanges data using UDP communication. The experimental apparatus is depicted in Figure 6.13. Each participant sat in front of the computer screen and grasped the handle of the haptic interface. They could not see or hear each other and were not allowed to talk. The subjects were instructed to perform reaching movements in the horizontal plane.

Software application

The devices can be used with dedicated libraries provided by Articares. The H-MAN's control library consists of two DLL files (Dynamic Liked Library):

1. Articares.Core.ArticaresComm: comprises 3 classes
    (a) Articares.Core.ArticaresComm: the H-MAN's commands can be found in the 'Articares.Core.ArticaresComm' class.
    (b) Articares.Core.HMANData: structure containing H-MAN's response parameters.
    (c) common.Helper: class containing helping functions.
2. NLog.dll: used to create an 'NLog' text file, in the 'C:' directory.

Our goal was to develop a dyadic set-up, in which the handles of the two H-Mans are mechanically coupled through a virtual spring. To this purpose, we developed a dedicated Matlab API which runs in each control computer which allows the communication between the two systems and manages the timing and the communication with the experiment which is developed in Simulink. We used two computers - one for each H-Man - because the H-Mans' IP addresses were the same hard-coded in the firmware and could not be changed externally. In each computer, two Matlab instances run in parallel. The first Matlab instance runs the Matlab API, which manages the communication with the device, and the second Matlab instance runs the Simulink model which manages the experiment design. The typical approach is to develop a stand-alone Simulink model. We developed a dedicated Matlab API to manage the communication with the device because the DLL files are not compatible with Simulink Desktop Real-Time.

The Matlab API, manages the communication and data exchange with the device itself, the Simulink model, and the other control computer. To have a real-time performances we used timer function each $tr = 0.005$ s (200 Hz):

- Create a “Articares.Core.ArticaresComm” class object and use this object in the next steps.
- Establish TCP connection with H-MAN; UDP connection with the Simulink model (experiment); UDP communication with the other control computer.
- Start exercise
- Call a timer function which
  - Reads positional data from another control computer;
  - Reads an enable message from the Simulink model which activates or deactivates the haptic rendering
  - Sends the control command to the H-Man using the SetTarget command.
  - Reads positional data from the H-Man
  - Sends them to the other control computer.
- Stop exercise
- Close TCP and UDP communication

The Simulink model uses the Simulink Desktop Real-Time toolbox. Simulink Desktop Real-Time is a self-targeting rapid prototyping system where the host and the target computer are the same computer. It provides a real-time kernel that allows to create and run hardware-in-the-loop simulations. The Simulink model runs at $tc = 0.001$ s (1000 Hz), and the refresh rates are different for the different components. The Simulink model:

- Exchanges data with the Matlab API (send the ENABLE, receives positional data each $tr = 0.005$ s, 200 Hz)
- Manages the experimental finite-state machine using the Stateflow toolbox (see Figure)
- Manages the visual feedback rendering (at $tv = 0.025$ s, 40 Hz), using Simulink 3D Animation (see Figure)
- Manages the data archiving (at $tv = 0.01$ s, 100 Hz)

The overall architecture is depicted in Figure 6.13

Visual Feedback

We developed a simple virtual environment where the elements of the reaching task are displayed so that the human subject has real-time visual feedback of their actions.

We used Simulink Virtual Reality Sink, an environment that allows to create *.wrl* extension files, i.e. virtual worlds realized in *VRML* (Virtual Reality Modeling Language).
In VRML, it is possible to create objects and assign them attributes (called 'nodes') - such as shape, dimension, and color - that can be modified through Simulink, in compliance with the real-time actions of the dyad. For example, the target object appears just when both the virtual partner's and human's cursors had been on the starting point for a whole second.

A fundamental feature of VRML is the points of view, i.e virtual cameras that 'shoot' the scene. We needed a 2D virtual space, so we placed the camera parallel to the frontal plane. Then, we wanted a world that was in scale with the virtual one without distortions, so we had to choose a small field of view (FOV). After setting $FOV = 0.045$ rad and knowing that the height of the display (x) was about 34 cm, we calculated the focal length (f) using the following equation:

$$\frac{x}{2f} = \tan\left(\frac{\text{FOV}}{2}\right) \tag{6.1}$$

with $0 < FOV < \pi$. The value that solves the equation is $f = 7.5$ m.

We then created all the objects needed depending on the applications.

Chapter 7

Conclusions and outlook

This chapter presents conclusions of the works conducted for this thesis and summarizes the major contributions. Moreover, we also discuss the study limitationa and potential future developments and applications.

7.1 Major contributions

7.1.1 The sense of agency in joint action

In Chapter 2 we addressed the subjective experience of control - i.e. the sense of agency - in interactive scenarios and how it influences and is influenced as the interaction unfolds. We proposed a computational model of joint action which attempts to interpret the subjective experience of agency in probabilistic (Bayesian) terms that are commonly used to address sensorimotor control. We extended the comparator model (Blakemore et al., 2002; Frith et al., 2000), commonly recognized in individual sensorimotor control and self-agency, to explicitly include a representation of the partner. We used game theoretic concepts to address the control level of joint action (Başar and Olsder, 1998). To be consistent with experimental reports of agency that we maintain different senses of self and joint agency (Bolt and Loehr, 2017; Bolt et al., 2016) we posited two separate observer-controllers pairs. The two loops continuously run in parallel assessing the hypothesis of acting alone or jointly with another agent. The likelihoods resulting from the comparison of expectations and sensory information determine the sense of self and joint agency which are used in the action selection process as weighting coefficients of the actions to be issued. This model provides a general formulation that reconciles the more mechanistic perspective of sensorimotor control and the more subjective experience of joint action. Further, the model can be used and easily extended to account for action observation and interactions with multiple agents.

7.1.2 From human-human interaction to human-artificial agent interaction

Computational models of joint action are pivotal as they allow to formulate predictions and formally describe humans' behavior and the underlying mechanisms. Mathematical accounts of the processes underpinnings of joint action lend themselves to be transferred to the development of biomimetic controllers which can be embedded in artificial agents. In Chapter 3 we developed a bio-inspired artificial partner architecture. The artificial partner consists of a simulated body and sensory system, a state observer - extended to explicitly develop and account for a partner model - and a game theoretic feedback controller. We tested the architecture with two experiments involving humans. We found that the artificial partner estimates human partner actions and that the goodness of the estimate depends on the coupling and the noise introduced. Further, the artificial partner is capable of learning to coordinate its motions with the ones of a human. Interestingly, we observed that humans interacting with the artificial partner assume similar behaviors to the ones assumed by humans interacting with other humans. Overall, we found that even a seemingly simple architecture based on a basic partner model and game theoretic controller is capable of interacting with a human in a biological manner. Here we presented and tested an architecture that can be used to further investigate joint action, by manipulating artificial partner features, but can also be used to develop novel approaches and protocols for motor and cognitive rehabilitation.

7.1.3 Negotiation through haptic communication

To address the multifaceted complexity of joint action computational models need to be informed by empirical observations. Different aspects of interaction need to be investigated with dedicated experiments. We addressed the partner model development and the emergence of coordination with a computational model of interaction (Chapter 4) and different experiments (Chapter 5, 6). As for the previous settings, we focused on scenarios where haptic constitutes the major communication channel.

The two via-points task

We were interested in investigating emergent coordination in a situation where multiple strategies exist and understanding the partner is not enough to converge to stable coordination. To this purpose, in the first set of experiments, presented in Chapter 5, we addressed a scenario in which players are required to achieve the same final target but have incompatible sub-goals. To establish a collaboration players had two cost-equivalent solutions and they needed to decide which one to pick by exchanging information through the haptic channel. We found that not all dyads converged successfully toward a collaboration. We found that the personal attitude toward the exploration may be the trigger to the exploration of coordination strategies and the emergence of stable coordination. These results provide insights into the understanding of how humans predict a partner's actions and the importance of combining such predictions with evidence to establish a coordination.

The multi via-points task

In another study, we were interested in investigating whether and to which extent prior knowledge affect the establishment of coordination. In two sets of experiments, presented in Chapter 6, we addressed whether coordination emerges when players are instructed to perform the same task and receive the same or different visual cues on possible coordination strategies. We predicted that humans being lazy by nature would lead the players to achieve a shared low-effort strategy. Interestingly, we observed that interacting agents did not achieve complete coordination - i.e. they did not completely align their motions in space and time. Further, we observed an instantaneous behavioral change as the coupling is activated. This is consistent with predictions made with the computational model of agency, where we found that as interaction cues are available, these are instantaneously incorporated and used in the sensorimotor control loop leading to the experience of a sense of joint agency and switch in the motor behavior to be issued.

7.2 General conclusion

In this thesis work, various aspects of interaction were investigated aiming at providing insights into the understanding of the sensorimotor mechanisms underlying joint action. Physical human-human interaction has been the focus of the current thesis, with a particular interest in emergent coordination in goal-directed motor action. We specifically focused on reaching tasks involving multiple via-points, manipulating the number and the configuration of intermediate points to be reached within a movement, the sensory information, and the prior cues provided to the interacting partners.

Joint action is a multifaceted and complex scenario where subjective and behavioral aspects are strongly intermingled, even more than in individual motor control, and cannot be addressed separately. Further, experiments and computational models of joint action inform and inspire each other. For this reason, we started by using a computational model of joint action, we then focused on a particular aspect of joint action developing a novel computational model and addressing decision-making through experiments. We then investigated more ecological movements where multiple possible solutions exist to achieve coordination.

The computational model of the sense of joint agency provides a tool to make predictions about the subjective experience of control in interactive scenarios. A future challenge should be the experimental study of the sense of joint agency to test model predictions. As previously mentioned the majority of previous work addressed agency using questionnaires or discrete action tasks. However, it would be useful to understand and evaluate the sense of agency in more ecological scenarios, using quantitative and objective metrics.

The artificial partner architecture constitutes a basic architecture for the development of intelligent agents. Here, a major effort has been put into the development of such architecture and in a brief test phase to demonstrate the working principles. However, more subjects should participate in a validation study and different experimental conditions should be tested (varying AP's knowledge about the human, or its attitude toward the interaction).

The presented dyadic experiments provide important empirical cues to the study of joint action. To adeptly compare the data in the two symmetric via-points task we should address other dyads in a PV (partner visible) condition. Finally, in the last experimental chapter, further testing may address different experimental conditions. For instance, different prior cues could be presented to the participant, or different sensory modalities coupling the two subjects could be tested to address how sensory modalities are combined to establish coordination.

This experiments-modeling iteration is useful by itself to fill the gaps in the understanding of the sensorimotor control and perceptual mechanisms underlying joint action. However, it may be transferred to the development of bio-inspired and bio-mimetic artificial tools which are every day more present in our daily life. Even though this thesis work does not explicitly address rehabilitation, one of the major long-term goal is to provide insights for a new era for neuromotor robotic rehabilitation - see Figure 7.1. We can use dyadic setups to understand the mechanisms underpinning therapist-patient interaction. But we can also conceive artificial therapists (AT) that improve and optimize the recovery of impaired patients.

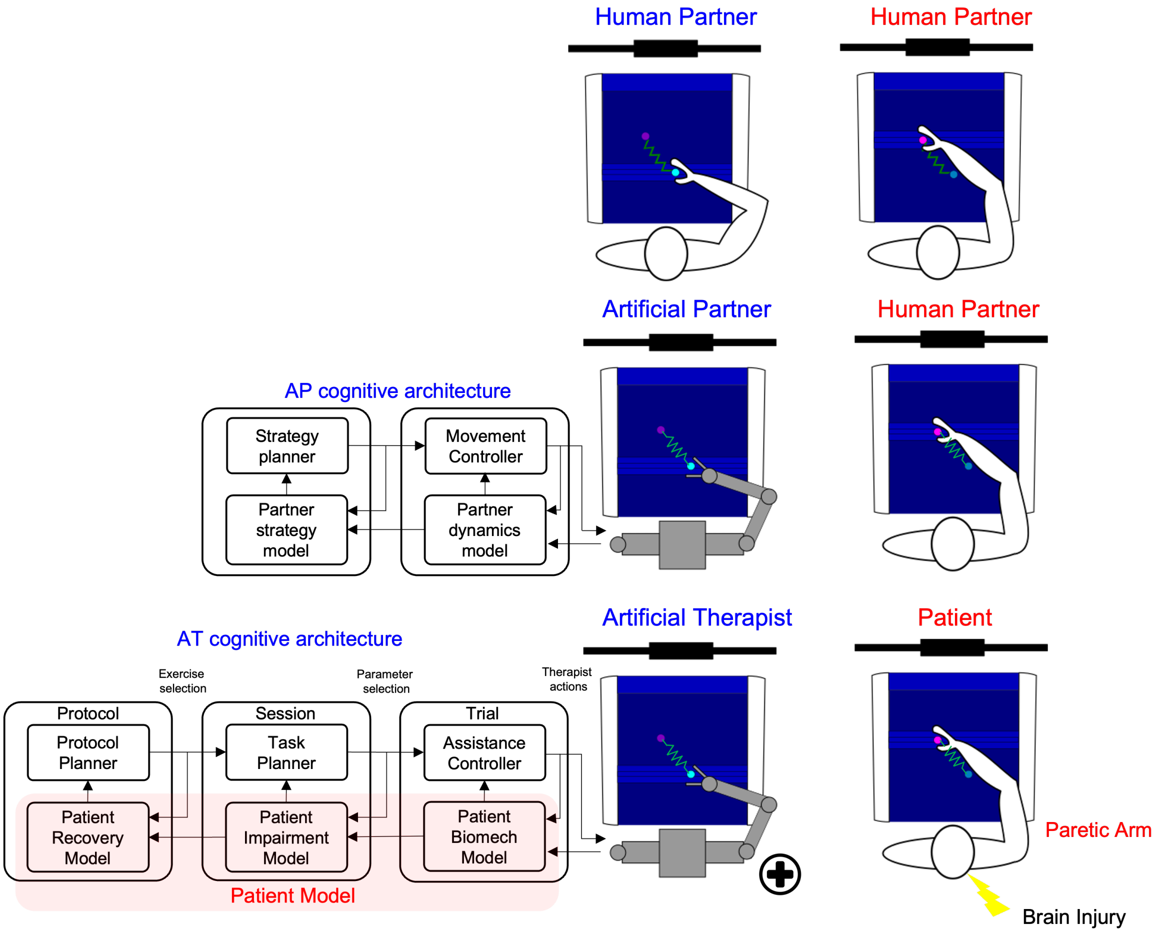

Figure 7.1 From human-human interaction to artificial therapists for neurorehabilitation.

7.3 List of Publications and Conferences

Journal articles

- De Vicariis, C., Chackochan, V. T., Sanguineti, V. (2022). Game theory and partner representation in joint action: toward a computational theory of joint agency. Phenomenology and the Cognitive Sciences, 1-30. DOI: 10.1007/s11097-022-09819-5
- De Vicariis, C., Pusceddu, G., Chackochan, V. T., Sanguineti, V. (2022). Artificial Partners to Understand Joint Action: Representing Others to Develop Effective Coordination. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 30, 1473-1482. DOI: 10.1109/TNSRE.2022.3176378

Book Chapters

- De Vicariis, C., Bandini, L., Chackochan, V. T., Sanguineti, V. Computational joint action: from emergent coordination to artificial partners. In Levin, Petrarca, Piscitelli, Summa (2023). Principles of motor control in clinical settings: from neuroscience to patient outcomes. Elsevier, in press.

Conference papers

- De Vicariis C., Chackochan V.T., Sanguineti V. (2020) The sense of joint agency in sensorimotor interactive task. Neuromatch Conference 3.0. (Online)
- De Vicariis C., Pusceddu G., Chackochan V.T., Sanguineti V. (2021) Artificial partners to investigate the development of optimal forms of collaboration in joint action. Neural Control of Movement (NCM) 19-21 April, 2021. (Online)
- De Vicariis C., Pusceddu G., Chackochan V.T., Sanguineti V. (2021) Artificial Partner to understand joint action and to make better robot therapy. International IEEE EMBS Conference on Neural Enigeering (NER) 4-6 May, 2021. (Online). Silver Poster Award.
- De Vicariis C., Sanguineti V. (2022) Partner representation, action selection and learning in joint coordination: experimental results and computational model, Neural Control of Movement (NCM) 25-29 July, 2022, Dublin.

- De Vicariis, C., Bandini, L., Chackochan, V. T., Sanguineti, V. (2023). Computational joint action: dynamical models to understand the development of joint coordination. Early version accepted at VIII Congress of the National Group of Bioengineering (GNB) 21-23 June, 2023, Padua.

- De Vicariis, C., Bandini, L., Pusceddu, G., Chackochan, V. T., Sanguineti, V. (2023). Artificial partners to understand joint coordination. Accepted at Joint Action Meeting (JAM) 10-12 July, 2023, Budapest.

- De Vicariis, C., Bandini, L., Chackochan, V. T., Sanguineti, V. (2023). Partner representation and decision-making in joint action. Accepted at Joint Action Meeting (JAM) 10-12 July, 2023, Budapest.

- De Vicariis, C., Bandini, L., Chackochan, V. T., Sanguineti, V. (2023). Partner representation and decision-making in joint action. Accepted at Joint Action Meeting (JAM) 10-12 July, 2023, Budapest.

References

Avila Mireles, E. J., Zenzeri, J., Squeri, V., Morasso, P., and De Santis, D. (2017). Skill learning and skill transfer mediated by cooperative haptic interaction. *IEEE Trans Neural Syst Rehabil Eng*, 25(7):832–843.

Bacharach, M. (1999). Interactive team reasoning: A contribution to the theory of co-operation. *Research in economics*, 53(2):117–147.

Başar, T. and Olsder, G. J. (1998). *Dynamic noncooperative game theory*. SIAM.

Beckers, N., van Asseldonk, E. H. F., and van der Kooij, H. (2020). Haptic human-human interaction does not improve individual visuomotor adaptation. *Sci Rep*, 10(1):19902.

Berger, U. (2007). Brown's original fictitious play. *Journal of Economic Theory*, 135(1):572–578.

Blakemore, S. J., Wolpert, D. M., and Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends Cogn Sci*, 6(6):237–242.

Bolt, N. K. and Loehr, J. D. (2017). The predictability of a partner's actions modulates the sense of joint agency. *Cognition*, 161:60–65.

Bolt, N. K., Poncelet, E. M., Schultz, B. G., and Loehr, J. D. (2016). Mutual coordination strengthens the sense of joint agency in cooperative joint action. *Conscious Cogn*, 46:173–187.

Brandi, M.-L., Kaifel, D., Bolis, D., and Schilbach, L. (2019). The interactive self–a review on simulating social interactions to understand the mechanisms of social agency. *i-com*, 18(1):17–31.

Braun, D. A., Ortega, P. A., and Wolpert, D. M. (2009). Nash equilibria in multi-agent motor interactions. *PLoS Comput Biol*, 5(8):e1000468.

Braun, D. A., Ortega, P. A., and Wolpert, D. M. (2011). Motor coordination: when two have to act as one. *Exp Brain Res*, 211(3-4):631–41.

Brown, G. W. (1951). Iterative solution of games by fictitious play. *Act. Anal. Prod Allocation*, 13(1):374.

Bruynseels, K., Santoni de Sio, F., and van den Hoven, J. (2018). Digital twins in health care: Ethical implications of an emerging engineering paradigm. *Front Genet*, 9:31.

Bryson, A. E. (1975). *Applied optimal control: optimization, estimation, and control.* Routledge.

Campolo, D., Tommasino, P., Gamage, K., Klein, J., Hughes, C. M. L., and Masia, L. (2014). H-man: a planar, h-shape cabled differential robotic manipulandum for experiments on human motor control. *J Neurosci Methods*, 235:285–97.

Carruthers, P. and Smith, P. K. (1996). *Theories of theories of mind.* Cambridge university press.

Casadio, M. and Sanguineti, V. (2012). Learning, retention, and slacking: a model of the dynamics of recovery in robot therapy. *IEEE Trans Neural Syst Rehabil Eng*, 20(3):286–96.

Casadio, M., Sanguineti, V., Morasso, P., and Solaro, C. (2008). Abnormal sensorimotor control, but intact force field adaptation, in multiple sclerosis subjects with no clinical disability. *Mult Scler*, 14(3):330–42.

Casadio, M., Sanguineti, V., Morasso, P. G., and Arrichiello, V. (2006). Braccio di ferro: a new haptic workstation for neuromotor rehabilitation. *Technol Health Care*, 14(3):123–42.

Chackochan, V. T. and Sanguineti, V. (2019). Incomplete information about the partner affects the development of collaborative strategies in joint action. *PLoS Comput Biol*, 15(12):e1006385.

Chambon, V., Sidarus, N., and Haggard, P. (2014). From action intentions to action effects: how does the sense of agency come about? *Front Hum Neurosci*, 8:320.

Cheng, S. and Sabes, P. N. (2006). Modeling sensorimotor learning with linear dynamical systems. *Neural Comput*, 18(4):760–93.

Clark, H. H. (1996). *Using language.* Cambridge university press.

Clark, H. H. (2020). Social actions, social commitments. In *Roots of human sociality*, pages 126–150. Routledge.

Claus, C. and Boutilier, C. (1998). The dynamics of reinforcement learning in cooperative multiagent systems. *AAAI/IAAI*, 1998(746-752):2.

Colomer, C., Dhamala, M., Ganesh, G., and Lagarde, J. (2022). Interacting humans use forces in specific frequencies to exchange information by touch. *Sci Rep*, 12(1):15752.

Curioni, A., Knoblich, G., Sebanz, N., Goswami, A., and Vadakkepat, P. (2019). Joint action in humans: A model for human-robot interactions. *Humanoid robotics: A reference*, pages 2149–2167.

De Vicariis, C., Chackochan, V. T., and Sanguineti, V. (2022). Game theory and partner representation in joint action: toward a computational theory of joint agency. *Phenomenology and the Cognitive Sciences.*

Dewey, J. A., Pacherie, E., and Knoblich, G. (2014). The phenomenology of controlling a moving object with another person. *Cognition*, 132(3):383–97.

Dumas, G., de Guzman, G. C., Tognoli, E., and Kelso, J. A. S. (2014a). The human dynamic clamp as a paradigm for social interaction. *Proc Natl Acad Sci U S A*, 111(35):E3726–34.

Dumas, G. and Fairhurst, M. T. (2021). Reciprocity and alignment: quantifying coupling in dynamic interactions. *R Soc Open Sci*, 8(5):210138.

Dumas, G., Kelso, J. A. S., and Nadel, J. (2014b). Tackling the social cognition paradox through multi-scale approaches. *Front Psychol*, 5:882.

Dumas, G., Moreau, Q., Tognoli, E., and Kelso, J. A. S. (2020). The human dynamic clamp reveals the fronto-parietal network linking real-time social coordination and cognition. *Cereb Cortex*, 30(5):3271–3285.

Emken, J. L., Benitez, R., Sideris, A., Bobrow, J. E., and Reinkensmeyer, D. J. (2007). Motor adaptation as a greedy optimization of error and effort. *J Neurophysiol*, 97(6):3997–4006.

Ernst, M. O. and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870):429–33.

Fletcher, P. C. and Frith, C. D. (2009). Perceiving is believing: a bayesian approach to explaining the positive symptoms of schizophrenia. *Nat Rev Neurosci*, 10(1):48–58.

Franklin, D. W. and Wolpert, D. M. (2011). Computational mechanisms of sensorimotor control. *Neuron*, 72(3):425–42.

Friston, K., Kilner, J., and Harrison, L. (2006). A free energy principle for the brain. *J Physiol Paris*, 100(1-3):70–87.

Frith, C. (2005). The self in action: lessons from delusions of control. *Conscious Cogn*, 14(4):752–70.

Frith, C. D., Blakemore, S. J., and Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. *Philos Trans R Soc Lond B Biol Sci*, 355(1404):1771–88.

Frith, C. D. and Frith, U. (2008). Implicit and explicit processes in social cognition. *Neuron*, 60(3):503–10.

Fudenberg, D. and Levine, D. K. (1998). *The theory of learning in games*, volume 2. MIT press.

Gallagher, S. (2008). Direct perception in the intersubjective context. *Conscious Cogn*, 17(2):535–43.

Gallese, V. and Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends Cogn Sci*, 2(12):493–501.

Gallivan, J. P., Stewart, B. M., Baugh, L. A., Wolpert, D. M., and Flanagan, J. R. (2017). Rapid automatic motor encoding of competing reach options. *Cell Rep*, 19(4):890–893.

Gallotti, M. and Frith, C. D. (2013). Social cognition in the we-mode. *Trends Cogn Sci*, 17(4):160–5.

Ganesh, G., Takagi, A., Osu, R., Yoshioka, T., Kawato, M., and Burdet, E. (2014). Two is better than one: physical interactions improve motor performance in humans. *Sci Rep*, 4:3824.

Gangopadhyay, N. and Miyahara, K. (2015). Perception and the problem of access to other minds. *Philosophical Psychology*, 28(5):695–714.

Gillijns, S. and De Moor, B. (2007). Unbiased minimum-variance input and state estimation for linear discrete-time systems. *Automatica*, 43(1):111–116.

Goeree, J. K. and Holt, C. A. (1999). Stochastic game theory: for playing games, not just for doing theory. *Proc Natl Acad Sci U S A*, 96(19):10564–7.

Grau-Moya, J., Hez, E., Pezzulo, G., and Braun, D. A. (2013). The effect of model uncertainty on cooperation in sensorimotor interactions. *J R Soc Interface*, 10(87):20130554.

Grau-Moya, J., Ortega, P. A., and Braun, D. A. (2012). Risk-sensitivity in bayesian sensorimotor integration. *PLoS Comput Biol*, 8(9):e1002698.

Gupta, N. and Mehra, R. (1974). Computational aspects of maximum likelihood estimation and reduction in sensitivity function calculations. *IEEE transactions on automatic control*, 19(6):774–783.

Haar, S. and Donchin, O. (2020). A revised computational neuroanatomy for motor control. *J Cogn Neurosci*, 32(10):1823–1836.

Hadjiosif, A. M., Krakauer, J. W., and Haith, A. M. (2021). Did we get sensorimotor adaptation wrong? implicit adaptation as direct policy updating rather than forward-model-based learning. *J Neurosci*, 41(12):2747–2761.

Haggard, P. (2005). Conscious intention and motor cognition. *Trends Cogn Sci*, 9(6):290–5.

Haggard, P., Clark, S., and Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nat Neurosci*, 5(4):382–5.

Haken, H., Kelso, J. A., and Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biol Cybern*, 51(5):347–56.

Harsanyi, J. C. (1967). Games with incomplete information played by “bayesian” players, i–iii part i. the basic model. *Management science*, 14(3):159–182.

Haruno, M., Wolpert, D. M., and Kawato, M. (2001). Mosaic model for sensorimotor learning and control. *Neural Comput*, 13(10):2201–20.

Haruno, M., Wolpert, D. M., and Kawato, M. (2003). Hierarchical mosaic for movement generation. In *International congress series*, volume 1250, pages 575–590. Elsevier.

Heald, J. B., Ingram, J. N., Flanagan, J. R., and Wolpert, D. M. (2018). Multiple motor memories are learned to control different points on a tool. *Nat Hum Behav*, 2(4):300–311.

Held, R. and Freedman, S. J. (1963). Plasticity in human sensorimotor control. *Science*, 142(3591):455–62.

Hwang, E. J., Smith, M. A., and Shadmehr, R. (2006). Dissociable effects of the implicit and explicit memory systems on learning control of reaching. *Exp Brain Res*, 173(3):425–37.

Izawa, J., Asai, T., and Imamizu, H. (2016). Computational motor control as a window to understanding schizophrenia. *Neurosci Res*, 104:44–51.

Jarrassé, N., Charalambous, T., and Burdet, E. (2012). A framework to describe, analyze and generate interactive motor behaviors. *PLoS One*, 7(11):e49945.

Kager, S., Hussain, A., Budhota, A., Dailey, W. D., Hughes, C. M., Deshmukh, V. A., Kuah, C. W., Ng, C. Y., Yam, L. H., Xiang, L., Ang, Jr, M. H., Chua, K. S., and Campolo, D. (2019a). Work with me, not for me: Relationship between robotic assistance and performance in subacute and chronic stroke patients. *J Rehabil Assist Technol Eng*, 6:2055668319881583.

Kager, S., Hussain, A., Cherpin, A., Melendez-Calderon, A., Takagi, A., Endo, S., Burdet, E., Hirche, S., Ang, M. H., and Campolo, D. (2019b). The effect of skill level matching in dyadic interaction on learning of a tracing task. *IEEE Int Conf Rehabil Robot*, 2019:824–829.

Kahl, S. and Kopp, S. (2018). A predictive processing model of perception and action for self-other distinction. *Front Psychol*, 9:2421.

Kapteyn, M. G., Pretorius, J. V., and Willcox, K. E. (2021). A probabilistic graphical model foundation for enabling predictive digital twins at scale. *Nature Computational Science*, 1(5):337–347.

Kelso, J. A. S., de Guzman, G. C., Reveley, C., and Tognoli, E. (2009). Virtual partner interaction (vpi): exploring novel behaviors via coordination dynamics. *PLoS One*, 4(6):e5749.

Knoblich, G., Butterfill, S., and Sebanz, N. (2011). Psychological research on joint action: theory and data. *Psychology of learning and motivation*, 54:59–101.

Knoblich, G. and Jordan, J. S. (2003). Action coordination in groups and individuals: learning anticipatory control. *J Exp Psychol Learn Mem Cogn*, 29(5):1006–16.

Kodl, J., Ganesh, G., and Burdet, E. (2011). The cns stochastically selects motor plan utilizing extrinsic and intrinsic representations. *PLoS One*, 6(9):e24229.

Körding, K. P. and Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, 427(6971):244–7.

Kourtis, D., Sebanz, N., and Knoblich, G. (2013). Predictive representation of other people's actions in joint action planning: an eeg study. *Soc Neurosci*, 8(1):31–42.

Kourtis, D., Woźniak, M., Sebanz, N., and Knoblich, G. (2019). Evidence for we-representations during joint action planning. *Neuropsychologia*, 131:73–83.

Krakauer, J. W., Ghez, C., and Ghilardi, M. F. (2005). Adaptation to visuomotor transformations: consolidation, interference, and forgetting. *J Neurosci*, 25(2):473–8.

Krakauer, J. W., Ghilardi, M. F., and Ghez, C. (1999). Independent learning of internal models for kinematic and dynamic control of reaching. *Nat Neurosci*, 2(11):1026–31.

Küçüktabak, E. B., Kim, S. J., Wen, Y., Lynch, K., and Pons, J. L. (2021). Human-machine-human interaction in motor control and rehabilitation: a review. *J Neuroeng Rehabil*, 18(1):183.

Lackner, J. R. and DiZio, P. (2005). Motor control and learning in altered dynamic environments. *Curr Opin Neurobiol*, 15(6):653–9.

Lafleur, A., Soulières, I., and Forgeot d'Arc, B. (2020). Sense of agency: Sensorimotor signals and social context are differentially weighed at implicit and explicit levels. *Conscious Cogn*, 84:103004.

Le Bars, S., Devaux, A., Nevidal, T., Chambon, V., and Pacherie, E. (2020). Agents' pivotality and reward fairness modulate sense of agency in cooperative joint action. *Cognition*, 195:104117.

Leibfried, F., Grau-Moya, J., and Braun, D. A. (2015). Signaling equilibria in sensorimotor interactions. *Cognition*, 141:73–86.

Li, Y., Carboni, G., Gonzalez, F., Campolo, D., and Burdet, E. (2019). Differential game theory for versatile physical human–robot interaction. *Nature Machine Intelligence*, 1(1):36–43.

Li, Y., Sena, A., Wang, Z., Xing, X., Babic, J., van Asseldonk, E. H., and Burdet, E. (2022). A review on interaction control for contact robots through intent detection. *Progress in Biomedical Engineering*.

Lindig-León, C., Schmid, G., and Braun, D. A. (2021a). Bounded rational response equilibria in human sensorimotor interactions. *Proc Biol Sci*, 288(1962):20212094.

Lindig-León, C., Schmid, G., and Braun, D. A. (2021b). Nash equilibria in human sensorimotor interactions explained by q-learning with intrinsic costs. *Sci Rep*, 11(1):20779.

Liu, D. and Todorov, E. (2007). Evidence for the flexible sensorimotor strategies predicted by optimal feedback control. *Journal of Neuroscience*, 27(35):9354–9368.

Loehr, J. D. (2018). Shared credit for shared success: Successful joint performance strengthens the sense of joint agency. *Conscious Cogn*, 66:79–90.

Loehr, J. D. and Vesper, C. (2016). The sound of you and me: Novices represent shared goals in joint action. *Q J Exp Psychol (Hove)*, 69(3):535–47.

Lokesh, R., Sullivan, S., Calalo, J. A., Roth, A., Swanik, B., Carter, M. J., and Cashaback, J. G. A. (2022). Humans utilize sensory evidence of others' intended action to make online decisions. *Sci Rep*, 12(1):8806.

Marchal-Crespo, L. and Reinkensmeyer, D. J. (2009). Review of control strategies for robotic movement training after neurologic injury. *Journal of neuroengineering and rehabilitation*, 6(1):1–15.

McKelvey, R. D. and Palfrey, T. R. (1995). Quantal response equilibria for normal form games. *Games and economic behavior*, 10(1):6–38.

McKelvey, R. D. and Palfrey, T. R. (1998). Quantal response equilibria for extensive form games. *Experimental economics*, 1:9–41.

Meinhardt-Injac, B., Daum, M. M., Meinhardt, G., and Persike, M. (2018). The two-systems account of theory of mind: Testing the links to social- perceptual and cognitive abilities. *Front Hum Neurosci*, 12:25.

Melendez-Calderon, A., Bagutti, L., Pedrono, B., and Burdet, E. (2011). Hi5: A versatile dual-wrist device to study human-human interaction and bimanual control. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2578–2583. IEEE.

Melendez-Calderon, A., Komisar, V., and Burdet, E. (2015). Interpersonal strategies for disturbance attenuation during a rhythmic joint motor action. *Physiol Behav*, 147:348–58.

Moore, J. and Haggard, P. (2008). Awareness of action: Inference and prediction. *Conscious Cogn*, 17(1):136–44.

Moore, J. W. and Fletcher, P. C. (2012). Sense of agency in health and disease: a review of cue integration approaches. *Conscious Cogn*, 21(1):59–68.

Nash, J. (1951). Non-Cooperative Games. *Annals of Mathematics*, 54(2):286–295.

Nash, J. (1953). Two-person cooperative games. *Econometrica: Journal of the Econometric Society*, pages 128–140.

Nash, J. F. et al. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, 36(1):48–49.

Newton, J. (2018). Evolutionary game theory: A renaissance. *Games*, 9(2):31.

Noy, L., Weiser, N., and Friedman, J. (2017). Synchrony in joint action is directed by each participant's motor control system. *Front Psychol*, 8:531.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the edinburgh inventory. *Neuropsychologia*, 9(1):97–113.

Oullier, O., de Guzman, G. C., Jantzen, K. J., Lagarde, J., and Kelso, J. A. S. (2008). Social coordination dynamics: measuring human bonding. *Soc Neurosci*, 3(2):178–92.

Pacherie, E. (2008). The phenomenology of action: a conceptual framework. *Cognition*, 107(1):179–217.

Pacherie, E. (2012). The phenomenology of joint action: Self-agency vs joint-agency. *Joint Attention: New Developments*, pages 343–389.

Pacherie, E. (2014). How does it feel to act together? *Phenomenology and the Cognitive Sciences*, 13(1):25–46.

Pesquita, A., Whitwell, R. L., and Enns, J. T. (2018). Predictive joint-action model: A hierarchical predictive approach to human cooperation. *Psychon Bull Rev*, 25(5):1751–1769.

Pezzulo, G., Donnarumma, F., and Dindo, H. (2013). Human sensorimotor communication: a theory of signaling in online social interactions. *PLoS One*, 8(11):e79876.

Pfeiffer, U. J., Schilbach, L., Jording, M., Timmermans, B., Bente, G., and Vogeley, K. (2012). Eyes on the mind: investigating the influence of gaze dynamics on the perception of others in real-time social interaction. *Front Psychol*, 3:537.

Prinz, A. A., Abbott, L. F., and Marder, E. (2004). The dynamic clamp comes of age. *Trends Neurosci*, 27(4):218–24.

Reed, K., Peshkin, M., Hartmann, M. J., Grabowecky, M., Patton, J., and Vishton, P. M. (2006). Haptically linked dyads: are two motor-control systems better than one? *Psychol Sci*, 17(5):365–6.

Reed, K. B. and Peshkin, M. A. (2008). Physical collaboration of human-human and human-robot teams. *IEEE Trans Haptics*, 1(2):108–120.

Rivera, L. F., Jiménez, M., Angara, P., Villegas, N. M., Tamura, G., and Müller, H. A. (2019). Towards continuous monitoring in personalized healthcare through digital twins. In *Proceedings of the 29th Annual International Conference on Computer Science and Software Engineering*, pages 329–335.

Rubinstein, A. (1998). *Modeling bounded rationality*. MIT press.

Sabu, S., Curioni, A., Vesper, C., Sebanz, N., and Knoblich, G. (2020). How does a partner's motor variability affect joint action? *PLoS One*, 15(10):e0241417.

Scheidt, R. A., Dingwell, J. B., and Mussa-Ivaldi, F. A. (2001). Learning to move amid uncertainty. *J Neurophysiol*, 86(2):971–85.

Schmidt, R. C. and O'Brien, B. (1997). Evaluating the dynamics of unintended interpersonal coordination. *Ecological Psychology*, 9(3):189–206.

Schmidt, R. C. and Richardson, M. J. (2008). Dynamics of interpersonal coordination. *Coordination: Neural, behavioral and social dynamics*, pages 281–308.

Sebanz, N., Bekkering, H., and Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends Cogn Sci*, 10(2):70–6.

Sebanz, N. and Knoblich, G. (2009). Prediction in joint action: what, when, and where. *Top Cogn Sci*, 1(2):353–67.

Sebanz, N., Knoblich, G., and Prinz, W. (2005). How two share a task: corepresenting stimulus-response mappings. *J Exp Psychol Hum Percept Perform*, 31(6):1234–46.

Semprini, M., Laffranchi, M., Sanguineti, V., Avanzino, L., De Icco, R., De Michieli, L., and Chiappalone, M. (2018). Technological approaches for neurorehabilitation: From robotic devices to brain stimulation and beyond. *Front Neurol*, 9:212.

Shadmehr, R. and Krakauer, J. W. (2008). A computational neuroanatomy for motor control. *Exp Brain Res*, 185(3):359–81.

Shadmehr, R. and Mussa-Ivaldi, F. A. (1994). Adaptive representation of dynamics during learning of a motor task. *J Neurosci*, 14(5 Pt 2):3208–24.

Sharp, A. A., O'Neil, M. B., Abbott, L. F., and Marder, E. (1993a). The dynamic clamp: artificial conductances in biological neurons. *Trends Neurosci*, 16(10):389–94.

Sharp, A. A., O'Neil, M. B., Abbott, L. F., and Marder, E. (1993b). Dynamic clamp: computer-generated conductances in real neurons. *J Neurophysiol*, 69(3):992–5.

Shumway, R. H. and Stoffer, D. S. (1982). An approach to time series smoothing and forecasting using the em algorithm. *Journal of time series analysis*, 3(4):253–264.

Silver, C. A., Tatler, B. W., Chakravarthi, R., and Timmermans, B. (2021). Social agency as a continuum. *Psychon Bull Rev*, 28(2):434–453.

Smith, M. A., Ghazizadeh, A., and Shadmehr, R. (2006). Interacting adaptive processes with different timescales underlie short-term motor learning. *PLoS Biol*, 4(6):e179.

Smith, M. A. and Shadmehr, R. (2005). Intact ability to learn internal models of arm dynamics in huntington's disease but not cerebellar degeneration. *J Neurophysiol*, 93(5):2809–21.

Stockinger, C., Focke, A., and Stein, T. (2014). Catch trials in force field learning influence adaptation and consolidation of human motor memory. *Front Hum Neurosci*, 8:231.

Synofzik, M., Vosgerau, G., and Newen, A. (2008). Beyond the comparator model: a multifactorial two-step account of agency. *Conscious Cogn*, 17(1):219–39.

Synofzik, M., Vosgerau, G., and Voss, M. (2013). The experience of agency: an interplay between prediction and postdiction. *Front Psychol*, 4:127.

Takagi, A., Beckers, N., and Burdet, E. (2016). Motion plan changes predictably in dyadic reaching. *PLoS One*, 11(12):e0167314.

Takagi, A., Ganesh, G., Yoshioka, T., Kawato, M., and Burdet, E. (2017). Physically interacting individuals estimate the partner's goal to enhance their movements. *Nature Human Behaviour*, 1(3):0054.

Takagi, A., Hirashima, M., Nozaki, D., and Burdet, E. (2019). Individuals physically interacting in a group rapidly coordinate their movement by estimating the collective goal. *Elife*, 8.

Takagi, A., Usai, F., Ganesh, G., Sanguineti, V., and Burdet, E. (2018). Haptic communication between humans is tuned by the hard or soft mechanics of interaction. *PLoS Comput Biol*, 14(3):e1005971.

Thoroughman, K. A. and Shadmehr, R. (2000). Learning of action through adaptive combination of motor primitives. *Nature*, 407(6805):742–7.

Todorov, E. (2004). Optimality principles in sensorimotor control. *Nat Neurosci*, 7(9):907–15.

Todorov, E., Erez, T., and Tassa, Y. (2012). Mujoco: A physics engine for model-based control. In *2012 IEEE/RSJ international conference on intelligent robots and systems*, pages 5026–5033. IEEE.

Todorov, E. and Jordan, M. I. (1998). Smoothness maximization along a predefined path accurately predicts the speed profiles of complex arm movements. *J Neurophysiol*, 80(2):696–714.

Todorov, E. and Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nat Neurosci*, 5(11):1226–35.

Tognoli, E., Zhang, M., Fuchs, A., Beetle, C., and Kelso, J. A. S. (2020). Coordination dynamics: A foundation for understanding social behavior. *Front Hum Neurosci*, 14:317.

Tuomela, R. (2007). *The philosophy of sociality: The shared point of view*. Oxford University Press.

Vallacher, R. R. and Wegner, D. M. (1989). Levels of personal agency: Individual variation in action identification. *Journal of Personality and Social psychology*, 57(4):660.

Vallacher, R. R., Wegner, D. M., and Somoza, M. P. (1989). That's easy for you to say: action identification and speech fluency. *J Pers Soc Psychol*, 56(2):199–208.

van der Wel, R. P. R. D. (2015). Me and we: Metacognition and performance evaluation of joint actions. *Cognition*, 140:49–59.

van der Wel, R. P. R. D., Knoblich, G., and Sebanz, N. (2011). Let the force be with us: dyads exploit haptic coupling for coordination. *J Exp Psychol Hum Percept Perform*, 37(5):1420–31.

Vesper, C., Abramova, E., Bütepage, J., Ciardo, F., Crossey, B., Effenberg, A., Hristova, D., Karlinsky, A., McEllin, L., Nijssen, S. R. R., Schmitz, L., and Wahn, B. (2016). Joint action: Mental representations, shared information and general mechanisms for coordinating with others. *Front Psychol*, 7:2039.

Vesper, C., Butterfill, S., Knoblich, G., and Sebanz, N. (2010). A minimal architecture for joint action. *Neural Netw*, 23(8-9):998–1003.

Vesper, C., Schmitz, L., and Knoblich, G. (2017). Modulating action duration to establish nonconventional communication. *J Exp Psychol Gen*, 146(12):1722–1737.

Vintsyuk, T. K. (1968). Speech discrimination by dynamic programming. *Cybernetics*, 4(1):52–57.

Von Neumann, J. and Morgenstern, O. (1944). Theory of games and economic behavior, science editions, 1964.

Wegner, D. M. (2003). The mind's best trick: how we experience conscious will. *Trends Cogn Sci*, 7(2):65–69.

Wegner, D. M. and Wheatley, T. (1999). Apparent mental causation. sources of the experience of will. *Am Psychol*, 54(7):480–92.

Wolpert, D. M. (1997). Computational approaches to motor control. *Trends Cogn Sci*, 1(6):209–16.

Wolpert, D. M., Doya, K., and Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philos Trans R Soc Lond B Biol Sci*, 358(1431):593–602.

Wolpert, D. M., Ghahramani, Z., and Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, 269(5232):1880–2.

Wolpert, D. M. and Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Netw*, 11(7-8):1317–29.

Wolpert, D. M. and Miall, R. C. (1996). Forward models for physiological motor control. *Neural Netw*, 9(8):1265–1279.

Yamaguchi, M., Wall, H. J., and Hommel, B. (2017). No evidence for shared representations of task sets in joint task switching. *Psychol Res*, 81(6):1166–1177.

Yoshida, W., Dolan, R. J., and Friston, K. J. (2008). Game theory of mind. *PLoS Comput Biol*, 4(12):e1000254.

Zalla, T. and Sperduti, M. (2015). The sense of agency in autism spectrum disorders: a dissociation between prospective and retrospective mechanisms? *Front Psychol*, 6:1278.

Zunino, A., Cavazza, J., Volpi, R., Morerio, P., Cavallo, A., Becchio, C., and Murino, V. (2020). Predicting intentions from motion: The subject-adversarial adaptation approach. *International Journal of Computer Vision*, 128:220–239.

Appendix A

Supplementary figures - Chapter 4

In this section, we report the model parameters derived from the identification procedure applied to the 2-via-points experiment as described in Chapter 4.

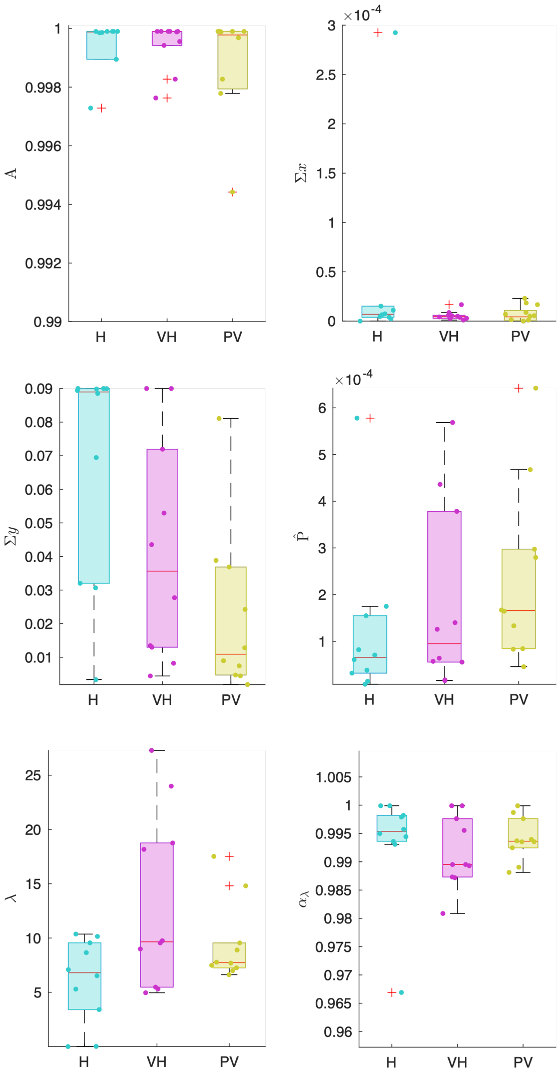

Figure A.1 Group differences in the estimated model parameters. Boxplots display median, 25th and 75th percentiles of the estimated parameters values, for H (light blue), VH (pink) and PV (yellow) groups.

Appendix B

Supplementary figures - Chapter 5

In this section, we report the trajectories of all the participants grouped into dyads for the two symmetrical via-points experiment presented in Chapter 5.

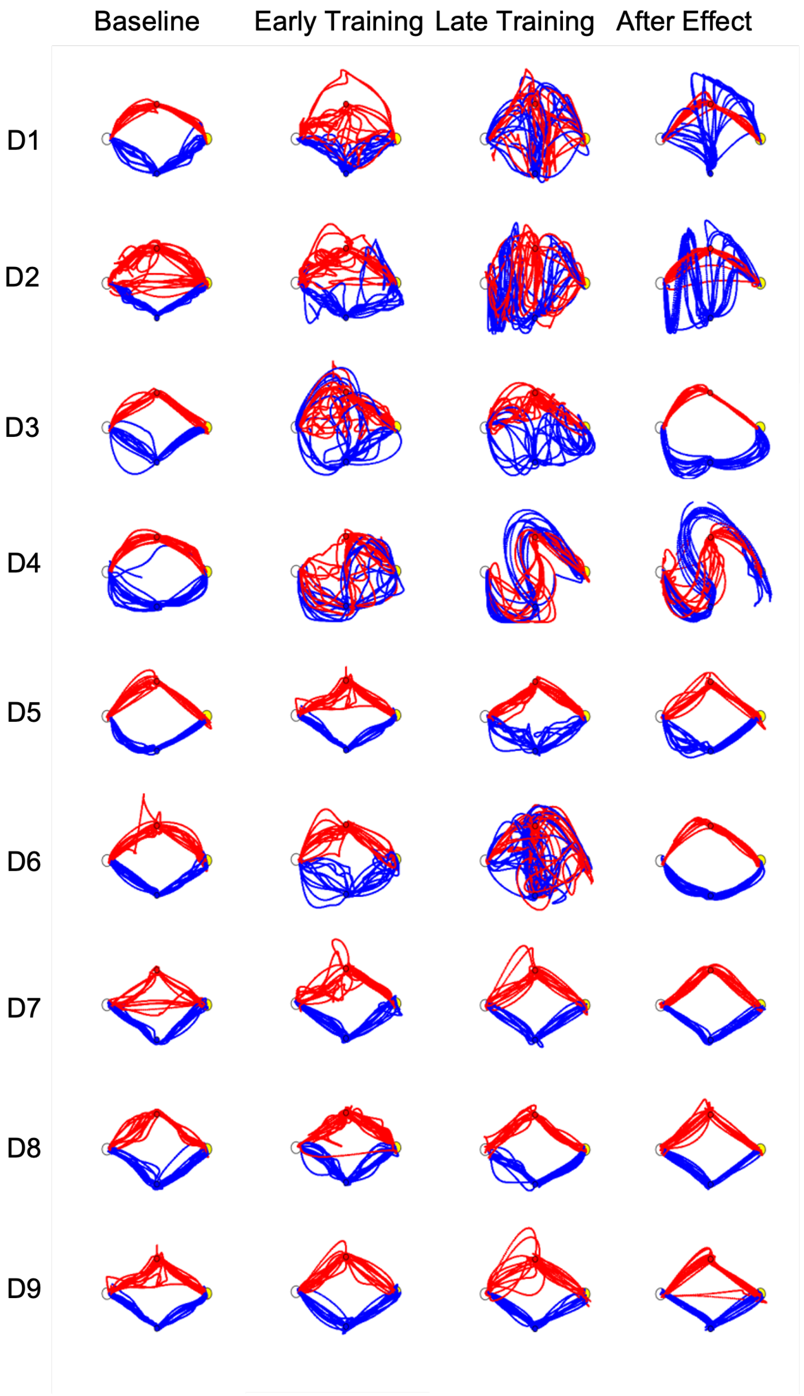

Figure B.1 Cursors trajectories are reported for all the dyads in the H condition in the baseline, early training, late training and after effect. In blue and red are reported trajectories for Player 1 and Player 2.

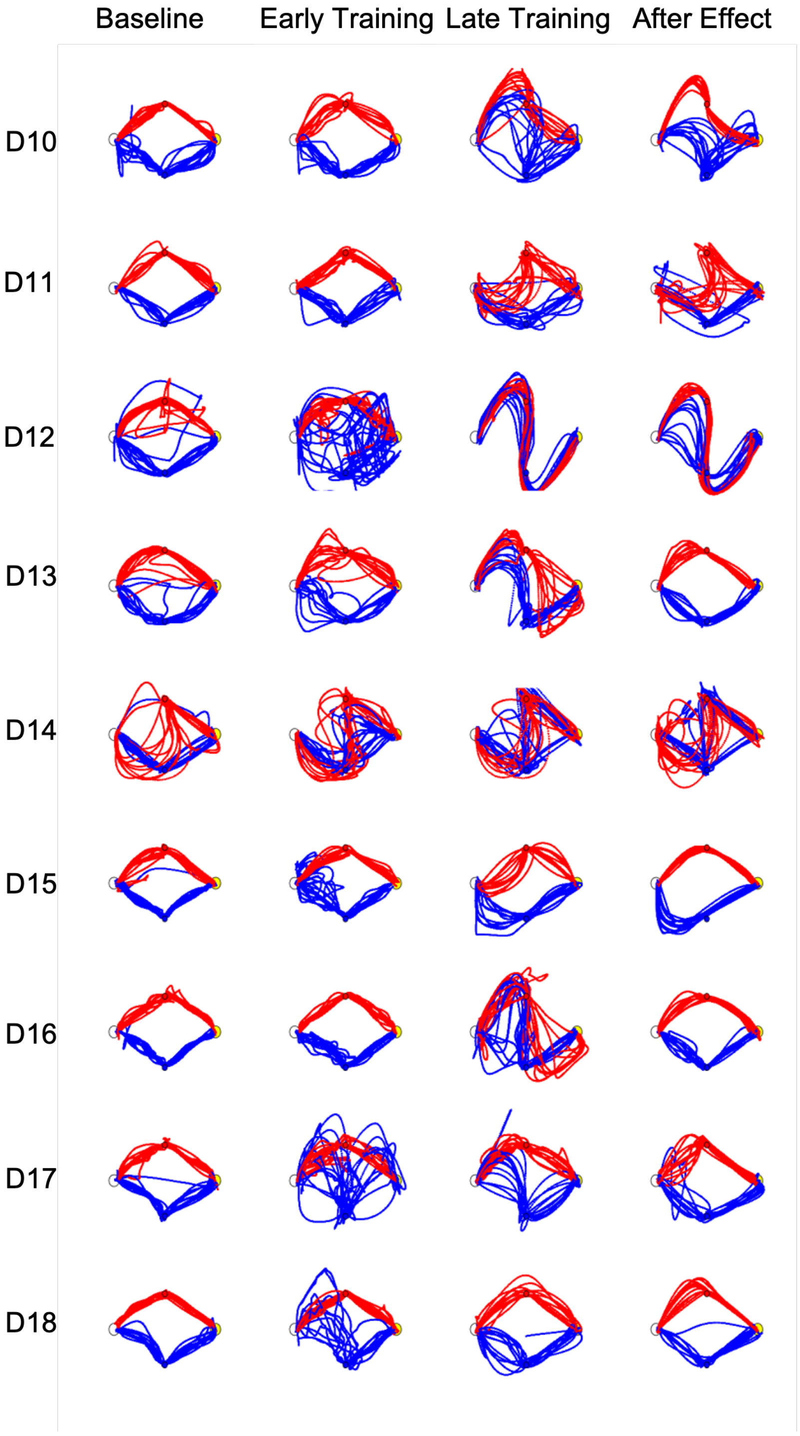

Figure B.2 Cursors trajectories are reported for all the dyads in the VH condition in the baseline, early training, late training and after effect. In blue and red are reported trajectories for Player 1 and Player 2.

Appendix C

Supplementary figures - Chapter 6

C.1 Experiment 1 - Collaboration

In this section, we report the trajectories of all the participants grouped into dyads for the collaboration experiment presented in Chapter 6.

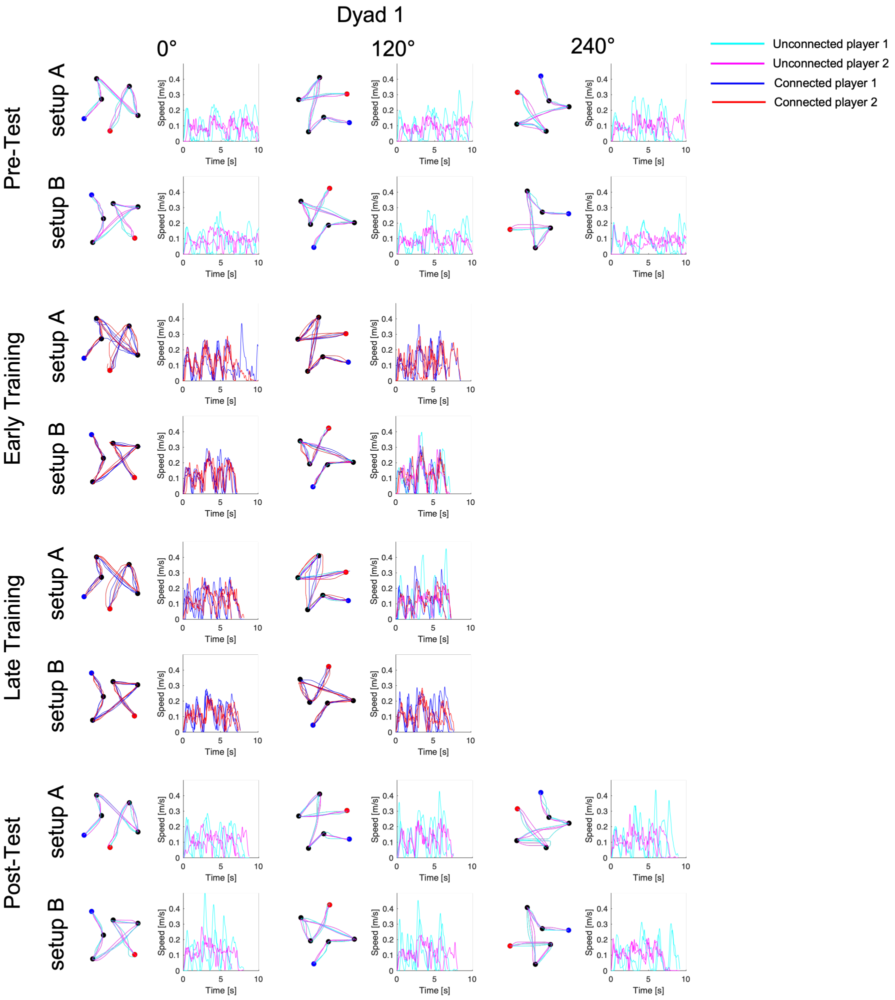

Figure C.1 Trajectories and speed profiles of dyad 1.In magenta and turquoise unconnected trials are reported, in red and blue trials with the active mechanical. Trajectories are reported for the two via-points configuration and for the three (in pre- and post-tests) and two (during training) rotations. Blue and turquoise are used to represent the trajectories and speed profiles of player 1 and red and magenta are used to represent player 2.

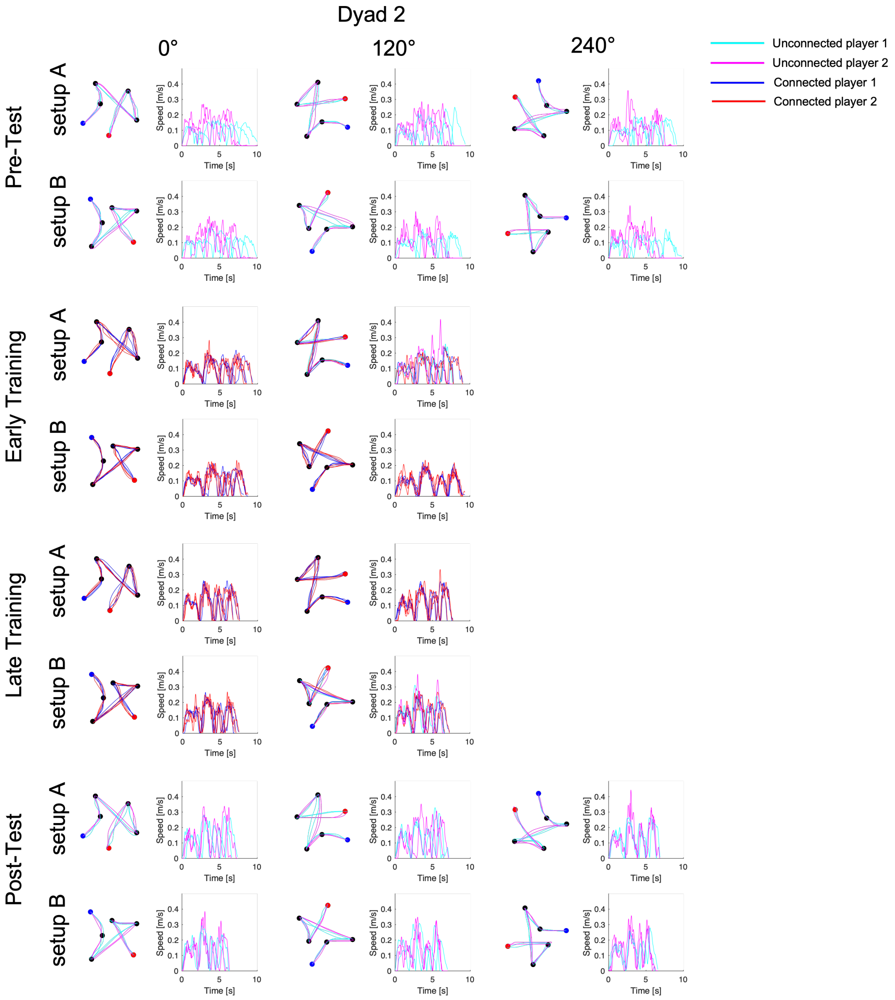

Figure C.2 Trajectories and speed profiles of dyad 2. In magenta and turquoise unconnected trials are reported, in red and blue trials with the active mechanical. Trajectories are reported for the two via-points configuration and for the three (in pre- and post-tests) and two (during training) rotations. Blue and turquoise are used to represent the trajectories and speed profiles of player 1 and red and magenta are used to represent player 2.

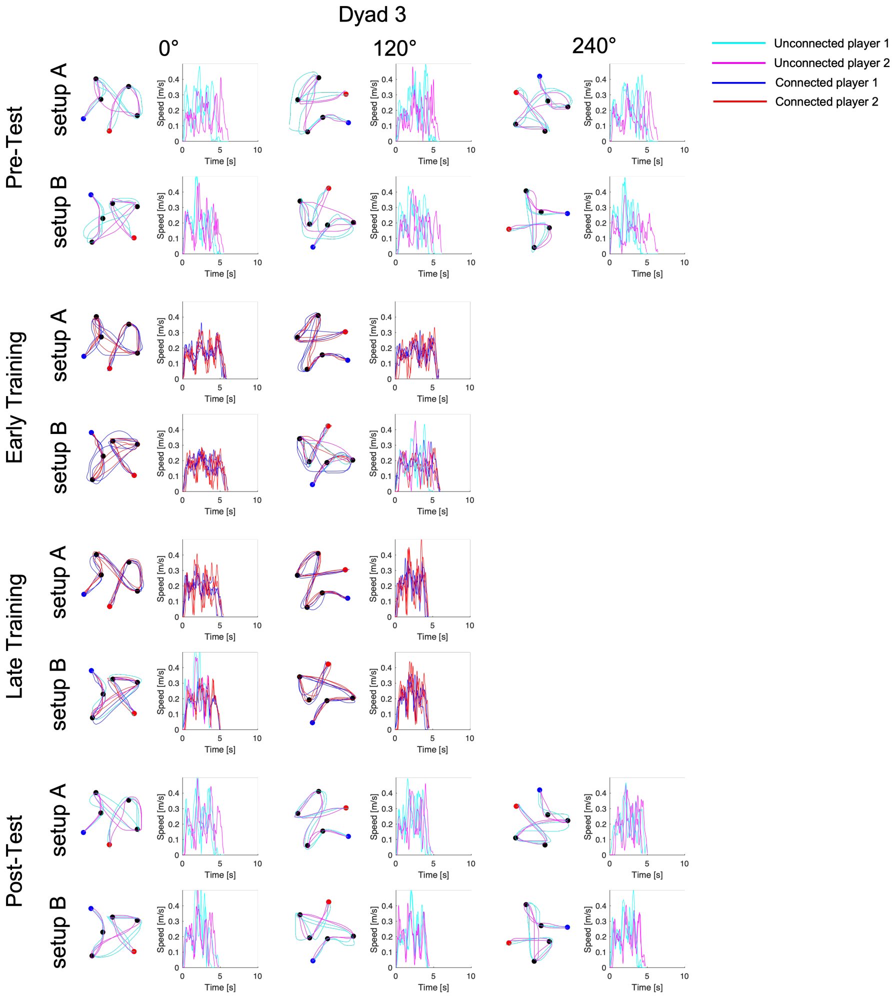

Figure C.3 Trajectories and speed profiles of dyad 3. In magenta and turquoise unconnected trials are reported, in red and blue trials with the active mechanical. Trajectories are reported for the two via-points configuration and for the three (in pre- and post-tests) and two (during training) rotations. Blue and turquoise are used to represent the trajectories and speed profiles of player 1 and red and magenta are used to represent player 2.

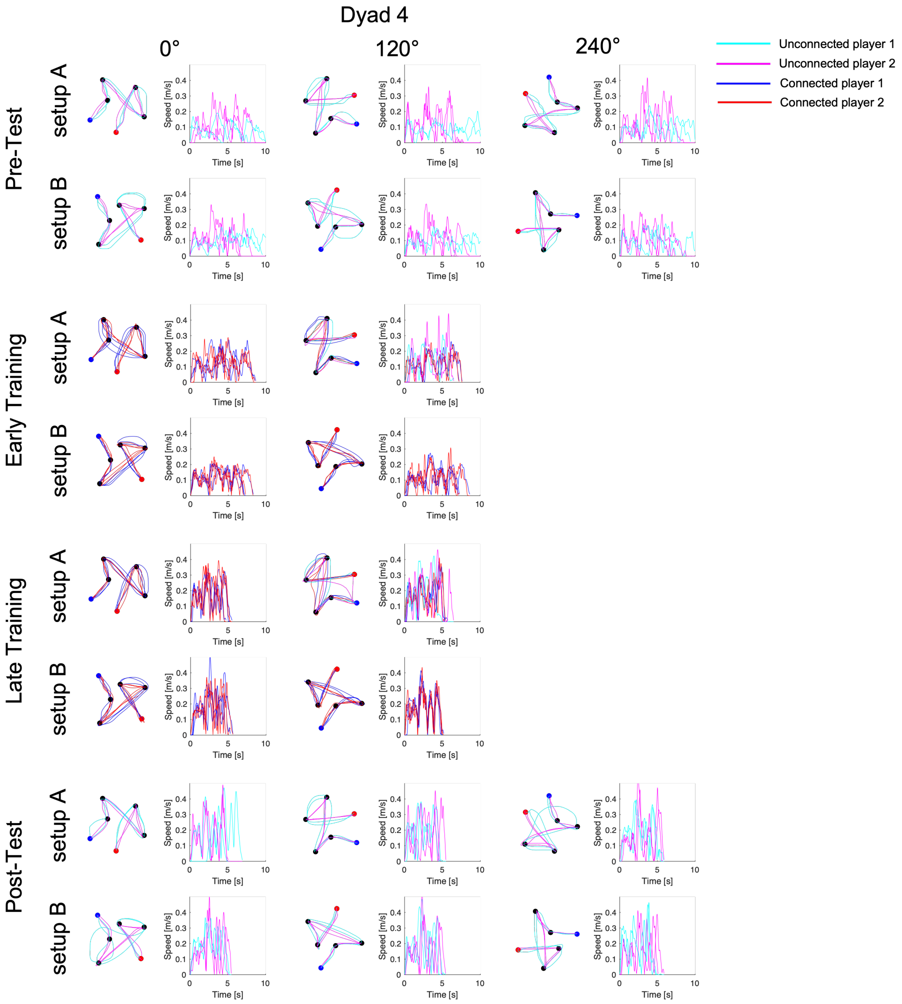

Figure C.4 Trajectories and speed profiles of dyad 4. In magenta and turquoise unconnected trials are reported, in red and blue trials with the active mechanical. Trajectories are reported for the two via-points configuration and for the three (in pre- and post-tests) and two (during training) rotations. Blue and turquoise are used to represent the trajectories and speed profiles of player 1 and red and magenta are used to represent player 2.

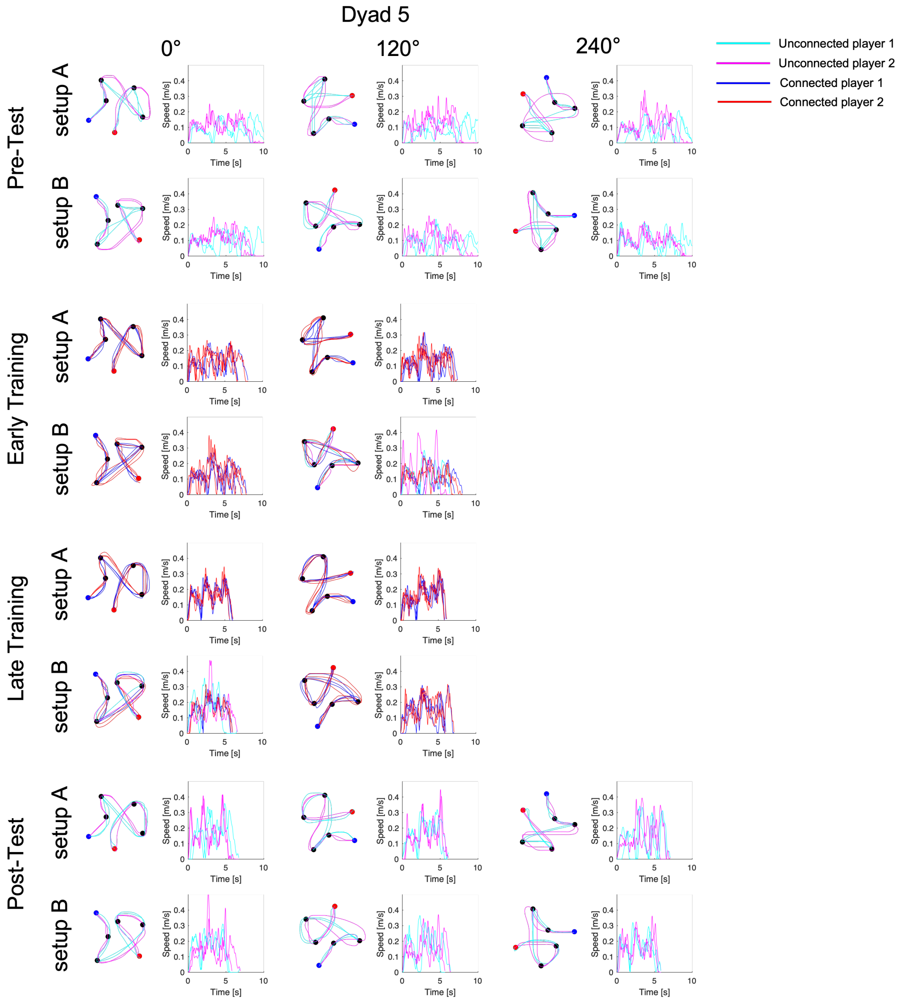

Figure C.5 Trajectories and speed profiles of dyad 5. In magenta and turquoise unconnected trials are reported, in red and blue trials with the active mechanical. Trajectories are reported for the two via-points configuration and for the three (in pre- and post-tests) and two (during training) rotations. Blue and turquoise are used to represent the trajectories and speed profiles of player 1 and red and magenta are used to represent player 2.

C.2 Experiment 2 - Negotiation

In this section, we report the trajectories of all the participants grouped into dyads for the negotiation experiment presented in Chapter 6.

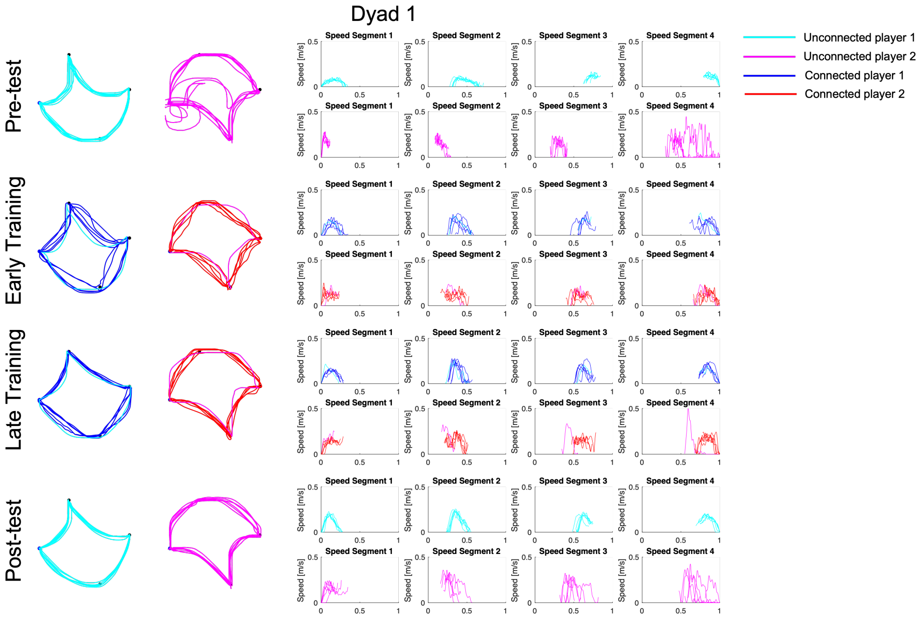

Figure C.6 Trajectories and speed profiles of dyad 1. In magenta and turquoise unconnected trials are reported, in red and blue trials with the active mechanical coupling. Speed profiles are divided by movement segment. Blue and turquoise are used to represent the trajectories and speed profiles of player 1 and red and magenta are used to represent player 2.

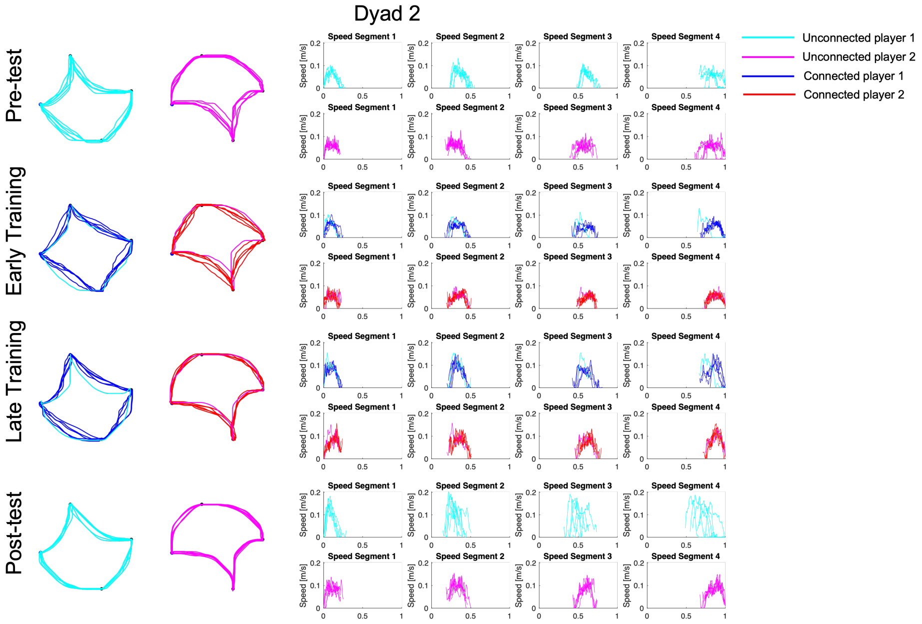

Figure C.7 Trajectories and speed profiles of dyad 2. In magenta and turquoise unconnected trials are reported, in red and blue trials with the active mechanical coupling. Speed profiles are divided by movement segment. Blue and turquoise are used to represent the trajectories and speed profiles of player 1 and red and magenta are used to represent player 2.

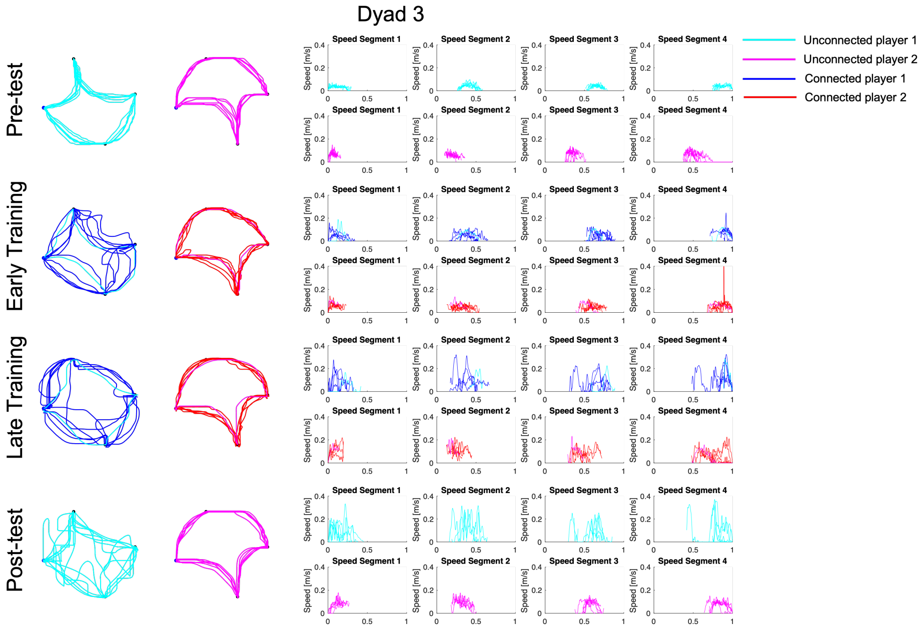

Figure C.8 Trajectories and speed profiles of dyad 3. In magenta and turquoise unconnected trials are reported, in red and blue trials with the active mechanical coupling. Speed profiles are divided by movement segment. Blue and turquoise are used to represent the trajectories and speed profiles of player 1 and red and magenta are used to represent player 2.

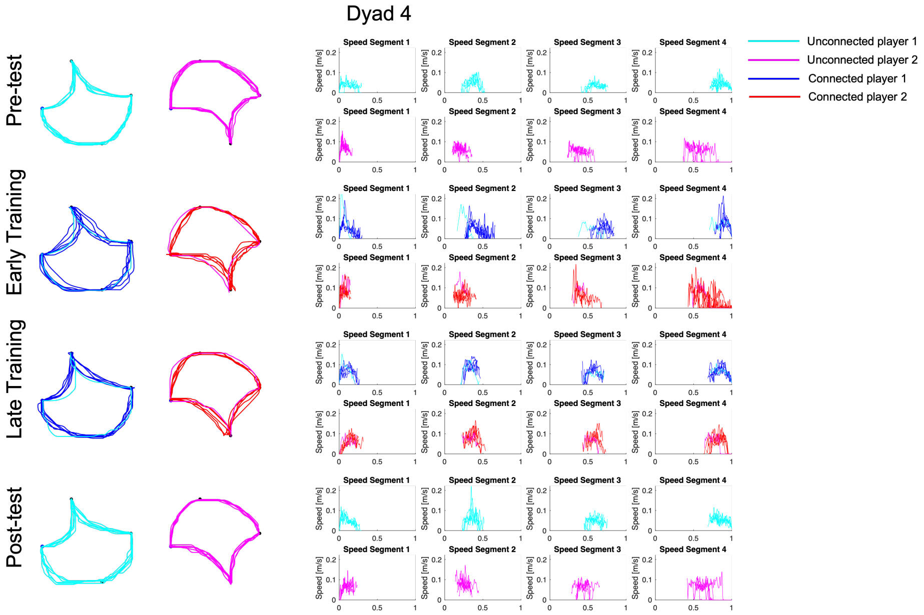

Figure C.9 Trajectories and speed profiles of dyad 4. In magenta and turquoise unconnected trials are reported, in red and blue trials with the active mechanical coupling. Speed profiles are divided by movement segment. Blue and turquoise are used to represent the trajectories and speed profiles of player 1 and red and magenta are used to represent player 2.

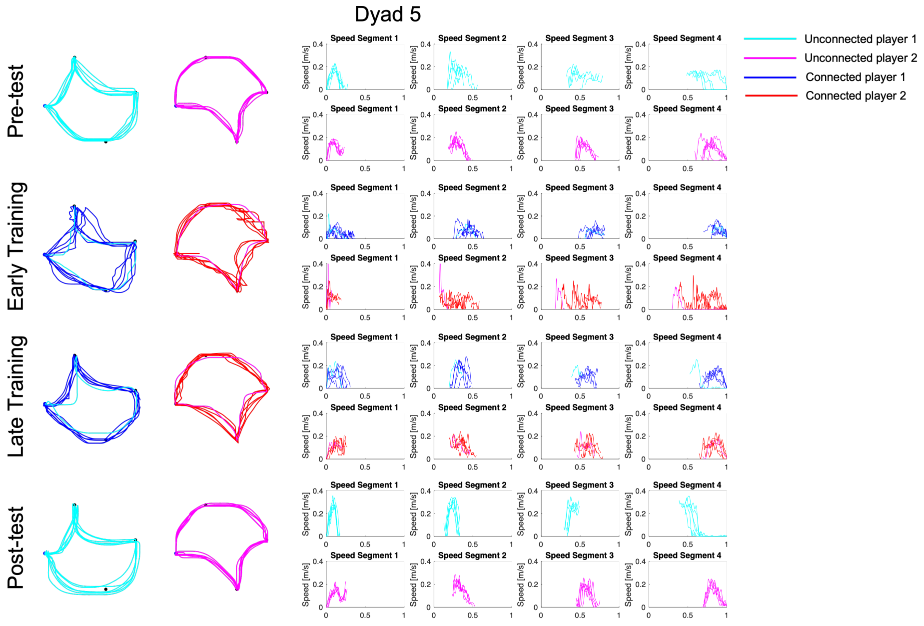

Figure C.10 Trajectories and speed profiles of dyad 5. In magenta and turquoise unconnected trials are reported, in red and blue trials with the active mechanical coupling. Speed profiles are divided by movement segment. Blue and turquoise are used to represent the trajectories and speed profiles of player 1 and red and magenta are used to represent player 2.

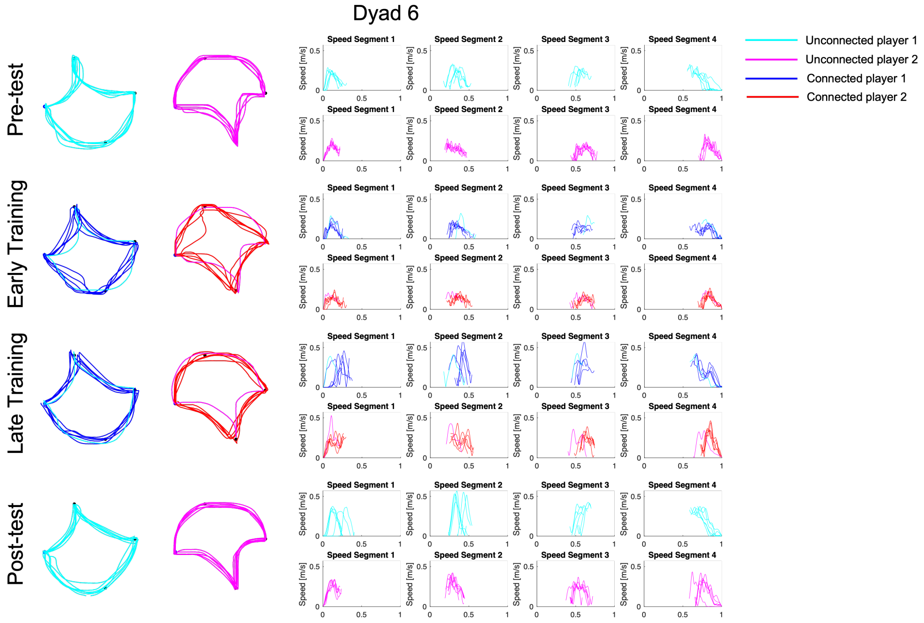

Figure C.11 Trajectories and speed profiles of dyad 6. In magenta and turquoise unconnected trials are reported, in red and blue trials with the active mechanical coupling. Speed profiles are divided by movement segment. Blue and turquoise are used to represent the trajectories and speed profiles of player 1 and red and magenta are used to represent player 2.

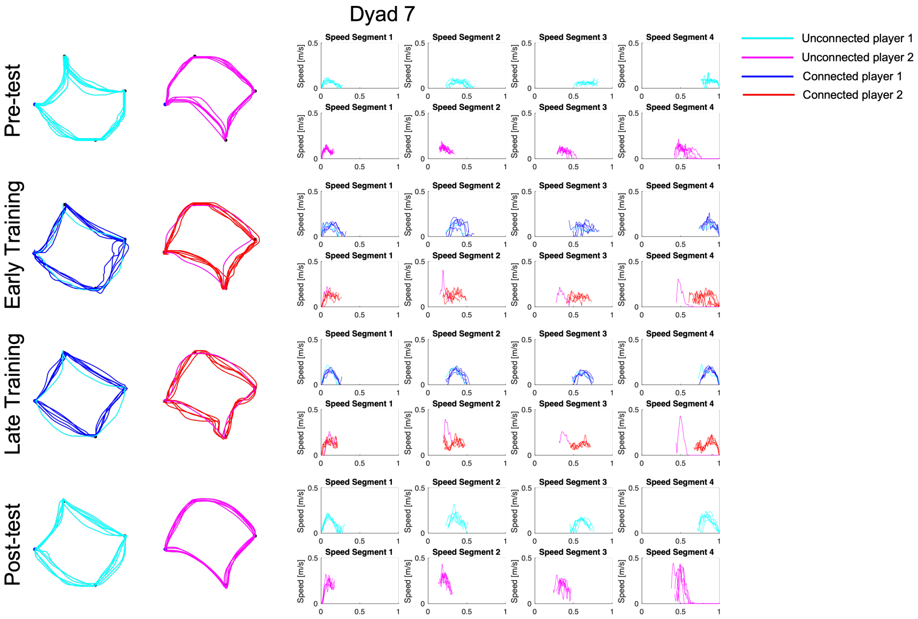

Figure C.12 Trajectories and speed profiles of dyad 7. In magenta and turquoise unconnected trials are reported, in red and blue trials with the active mechanical coupling. Speed profiles are divided by movement segment. Blue and turquoise are used to represent the trajectories and speed profiles of player 1 and red and magenta are used to represent player 2.