UNIVERSITY OF GENOVA

PHD PROGRAM IN BIOENGINEERING AND ROBOTICS

Neural bases and behavioural responses of *vitality forms* in human-human and human-robot interactions

by

Giada Lombardi

Thesis submitted for the degree of *Doctor of Philosophy* (36° cycle)

April 2024

Dr. Alessandra Sciutti	Supervisor
Dr. Giuseppe Di Cesare	Supervisor
Prof. Paolo Massobrio	Head of the PhD program

Thesis Jury:

Prof. Emily Cross, *University of Glasgow*	External examiner
Prof. Shelly Levy-Tzedek, *Ben-Gurion University of the Negev*	External examiner
Prof. Maura Casadio, *University of Genoa*	Internal examiner

Dibris

Department of Informatics, Bioengineering, Robotics and Systems Engineering

With all my love, I would like to dedicate this thesis to my parents,
for always supporting me, also in the most challenging times.
You have been and will be forever my guiding stars.

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.

Giada Lombardi
April 2024

Abstract

The observation of actions performed by others allows us to understand what they are doing and why. Such ability is related to the existence of a set of neurons, "mirror neurons", that discharge both when individuals perform a goal-directed action and when they observe another person performing the same action. However, besides action goal and intention, there is a third fundamental component that an observer may capture when viewing an action: its form. Indeed, actions can be executed with different forms, such as gently or rudely, vigorously or hesitantly, communicating the affective state of the agent as well as positive and negative attitudes towards the receiver. Daniel Stern defined these aspects of social communication “*vitality forms*”. The recognition and expression of vitality forms is already present in infants during mother-child interactions, suggesting that they represent a primordial way to relate to and understand others. Despite their pervasiveness in human life, research has always focused on the action goal, neglecting the role of vitality forms in the study of human behaviour and social interactions. In this view, my PhD project involves an interdisciplinary approach that spans the fields of neuroscience and robotics aiming to contribute to the enrichment of vitality forms research. Particularly, it deepens the study of neural correlates and behavioural responses underlying vitality forms processing in humans and proposes vitality forms as a novel and valuable tool to promote human—robot interactions.

The present thesis aimed to address the following objectives: 1) to deeply investigate in humans the activity of brain areas involved in vitality forms processing, 2) investigate the influence of vitality forms on action perception and expression during human—human interactions and 3) human—robot interactions. These challenges are outlined through three main studies, whose findings represent the contribution of this work.

Results of the first study show that both the dorso-central insula (DCI) and middle cingulate cortex (MCC) are endowed with a mirror mechanism specific for vitality forms processing. Indeed, a voxel based analysis revealed a significant correlation of the BOLD signal during observation and execution of actions endowed with vitality forms. Moreover, a multifiber tractography analysis showed that these two areas are anatomically connected, forming a circuit selective for the processing of hand-arm actions conveying vitality forms. This circuit may allow individuals to understand vitality forms expressed by others and to prepare an adequate motor response. This suggestion is supported by results provided in the second study, which show an influence of vitality forms conveyed by an agent, through different modalities, on the action perception (estimated action duration) and execution (kinematic parameters) of a receiver. Results of the third study assess that vitality forms expressed by a humanoid agent (the iCub robot), can induce the same influence on actions performed by participants. Most importantly, the iCub robot was used to demonstrate how positive (happy) and negative (angry) facial expressions can modify the perception of its action vitality forms and consequently modulate the motor response of the human partner.

Summing up, this thesis provides new neurophysiological and behavioural insights of vitality forms, highlighting their essential role in human—human interactions and proposing them as constitutive feature also for robots to reach an effective communication with humans. Results and methodology of this PhD project lay important foundations for future studies aiming to investigate vitality forms processing in clinical populations with social and motor impairments.

Table of contents

List of figures

Chapter 1

Introduction

"We naturally experience people in terms of their vitality. We intuitively evaluate their emotions, states of mind, what they are thinking and what they really mean, their authenticity, what they are likely to do next, as well as their health and illness on the basis of the vitality expressed in their almost constant movements."

(Stern, 2010)

1.1 Motivation

From birth humans interact with the world. Even before learning to speak, infants can communicate their needs and as soon as they learn how to perform simple actions, such as reaching or grasping objects, they immediately start to understand the consequences of other individuals performing the same actions. This innate non-verbal communication permeates social interactions also during adulthood. Indeed, people show a natural ability at interacting with each other, displaying behaviours that go beyond conscious control. For a clearer insight into this concept, let's think for a moment about our everyday life:

We are at a café and we observe someone moving the hand towards a cup. By observing this action, we can intuitively understand what the agent is doing (e.g. grasping the cup) and also why they are doing it (e.g. grasping the cup for drinking or grasping the cup for passing it towards another person sitting at the table).

From a neurophysiological perspective, there is evidence that this ability is related to the existence of a class of neurons (mirror neurons), that discharge both during action observation and action execution [Di Pellegrino et al. (1992); Gallese et al. (1996); Rizzolatti et al. (1996)]. Thanks to their activity, the motor processes that are primarily recruited in generating an action are also involved when someone observes the same action performed by others [Rizzolatti and Sinigaglia (2016)]. In this way, the observer is able to easily capture the action goal and intention of the agent.

However, action understanding can be further deepened by capturing another fundamental component: *how* the action is performed.

Following the previous example, grasping the cup could be executed vigorously or delicately and passing it rudely or gently. These subtle aspects of actions have been named *vitality forms* by Daniel Stern [Stern (2010)]. His intense research on infants has largely shown that vitality forms are a fundamental aspect of early interactions between children and their caregivers, thus representing a primordial way to relate to others.

Vitality forms shape our actions representing an externalization of our internal states, which are captured by others when socially interacting with them. Thus, vitality forms play a dual role in social communication: on the one hand, their execution enables an agent to express their attitudes and affective states. On the other hand, their perception enables a receiver to gather more information about the agent's state and consequently modulate their behaviour [Di Cesare et al. (2017)]. Without vitality forms our experience of others would be devoid of any affective colour and dynamics and we would be able to understand people only superficially [Di Cesare et al. (2020a)]. Nevertheless, despite their pervasiveness in human life and their crucial role in interpersonal relations, vitality forms are still poorly studied.

Sitting in the context of mirror neurons research, first experiments on the topic aimed to investigate what brain areas are active, together with the parieto-frontal network responsible of the action goal encoding, during vitality forms processing. First fMRI studies showed the selective activation of the dorso-central insula (DCI) during both the observation and execution of actions conveying vitality forms, suggesting the presence of mirror properties specific for vitality forms encoding [Di Cesare et al. (2020a)]. However, no studies have deeply investigated this point. Additionally, besides anatomical evidences, how DCI and parieto-frontal regions exchange information (in terms of causal influence, i.e. effective connectivity) during action observation and execution is still unknown, although it represents a fundamental issue to be addressed for a whole comprehension on how humans process actions in their everyday interactions.

Once defined what brain mechanisms occur during action forms processing, a step forward needs to be accomplished: what is the role and effect of vitality forms in social interactions? Particularly, it has been recently proposed that perceiving different vitality forms of an agent could influence the kinematic features characterizing actions performed by a receiver [Di Cesare et al. (2017)]. It is clear, therefore, that vitality forms enhance the quality of

interactions, by creating an unconscious and natural reciprocal comprehension between individuals together with a significant modulation of their motor behaviours.
Starting from this point, an intriguing issue concerns the possibility to investigate whether and how a humanoid agent able to generate vitality forms may induce a similar influence on humans. Giving robots, especially humanoid robots, the ability to express vitality forms would present some advantages in several situations, from basic daily communication with humans to more complex scenarios. For example, in elderly home care and education, robots could be more gentle, displaying slow and fluid motion, while in law enforcement and security, they could be able to communicate imperative commands and quickly convince a person to follow their instructions.
The present thesis aims to contribute to the enrichment of vitality forms research by deeply investigating their neural bases and behavioural responses in humans and propose them as pivotal feature to endow robots with in order to enrich social communication with the human partners. The core strength of this thesis lies in the interdisciplinary approach taken to explore vitality forms from different perspectives. The integration of several methodologies has allowed for a broader and more comprehensive understanding of the topic, establishing a deep connection between different fields of research such as neuroscience and robotics.

1.2 Research Objectives

Given the rationale behind this work, this section presents the main research objectives that motivated this thesis and shapes methodological approaches to address them.

- **RO1**: **Investigate the activity of brain areas involved in vitality forms processing in humans**

To address RO1, different neuroimaging techniques including functional magnetic resonance (fMRI), Diffusion Tensor Imaging (DTI) and Dynamic Causal Modelling (DCM) were used. Firstly, after an implementation of the experimental paradigm used in Di Cesare et al. (2015), a main fMRI experiment was carried out to assess brain areas involved in the observation and execution of actions expressing vitality forms. Secondly, DTI acquisition enabled to apply a tractography analysis and identify anatomical tracts connecting areas of interest. Finally, DCM was applied to functional data for the investigation of effective connectivity, i.e. direct causal influence between areas of interest.

- **RO2: Investigate the influence of vitality forms on action perception and expression during human—human interactions**

To address RO2, an experimental paradigm used in a previous kinematic study [Di Cesare et al. (2017)] was implemented by:
1) adding a new modality to convey vitality forms: physical contact. In particular, the robotic platform "Braccio di ferro" was used to deliver rude or gentle physical stimulations on participants right arm. The same robotic manipulandum was employed to record kinematic features of participants motor response.
2) adding a second cognitive task (action estimation task) to the existing motor task (action execution task). This enabled to investigate the influence of vitality forms expressed by an agent through different modalities both on the action execution and perception of the receiver.

- **RO3: Investigate the influence of vitality forms on action perception and expression during human—robot interactions**

This third objective aimed to answer a specific question: how additionally social cues such as facial expressions could modify the perception of action vitality forms expressed by an agent and consequently modify the motor response of a (human) receiver?
The iCub robot represented an ideal agent to assess this issue. By leveraging on the iCub robot, positive and negative facial expressions were easily associated to positive and negative action vitality forms obtaining four experimental conditions (two "congruent" and two "incongruent"). In my thesis, two strategies have been used to evaluate how actions performed by the iCub robot with gentle and rude vitality forms (already validated in previous experiments [Vannucci et al. (2018)]) were perceived when associated to happy and angry facial expressions:
1) measuring how participants modulate the kinematic features of their hand-arm actions in response to the iCub robot (kinematic experiment)
2) conducting questionnaires before and after the main kinematic study (behavioural studies).

These research challenges are outlined through four main experiments carried out during my PhD, whose findings are presented from Chapter 3 to Chapter 6. Each chapter presents a single study published in a related journal article. While the introduction and discussion of each chapter are built to link one research objective to the other, the sections of methods and results are, besides some light modifications (more detail in the thesis), do not deviate from the corresponding published article.

1.3 Thesis structure

In the present chapter, **Chapter 1**, I provide the rationale underlying my PhD project, the research objectives I planned to achieve and an overview of the thesis structure.

In **Chapter 2**, I examine the background in which my thesis sits and the current state of the art in the related field. Firstly, I focus on reviewing literature concerning the neurophysiological bases of action processing, with a specific attention to mirror neurons and action observation-execution network. Secondly, I present Daniel Stern’s definition of vitality forms. Finally, I describe methodology and results of first experiments on vitality forms in both neuroscience and robotics.

In **Chapter 3** and **Chapter 4**, I address RO1. Firstly, Chapter 3 provides results of an fMRI study, carried out at the Department of Neuroscience of Parma, concerning the presence of a circuit selectively activated during observation and execution of actions endowed with vitality forms. Secondly, starting from data presented in Chapter 3, in Chapter 4 I present a Dynamic Causal Modelling study carried out during my visiting research period at the Wellcome Centre for Human Neuroimaging of the University College of London (UCL).

In **Chapter 5**, I address RO2. Particularly, I describe methodology and results of an experiment carried out at the Italian Institute of Technology with the robotic platform "Braccio di Ferro". Furthermore, at the end of the chapter, Appendix A provides additional findings obtained in a preliminary behavioural study.

In **Chapter 6**, I address RO3. More specifically, Chapter 6 provides findings of a kinematic experiment carried out at the Italian Institute of Technology aiming to assess whether and how facial expressions of the humanoid robot iCub could modify the perception of its action vitality forms and consequently the motor response of participants. Furthermore, at the end of the chapter, Appendix B describes additional results obtained in two behavioural studies carried out before and after the kinematic experiment for the validation of experimental visual stimuli and to understand participants attitude towards the iCub robot respectively.

Finally, **Chapter 7** concludes the thesis by summarizing the achieved research objectives and by outlining future research directions based on the progress made in this work. Additionally, it includes a list of publications and conferences attended during my PhD to disseminate my research.

Chapter 2

Background

2.1 The Mirror Mechanism

2.1.1 The neuroanatomy of the mirror neuron system

The Mirror Mechanism is a brain mechanism that transforms sensory representations of others' behaviours into one's own motor or visceromotor representations concerning that behaviours [Rizzolatti and Sinigaglia (2016)]. This mechanism is related to the activity of a set of neurons, i.e. *mirror neurons*, that discharge both when individuals perform a motor act and when individuals observe another person performing a motor act with similar goal. Thus, the observation of an action produces a motor activation, as if the observer were actually programming the execution of the observed action [Ferrari and Rizzolatti (2015)].

Mirror neurons were first discovered in the ventral premotor cortex of the macaque monkey (area F5) [Di Pellegrino et al. (1992); Gallese et al. (1996); Rizzolatti et al. (1996)].

Electrophysiological studies showed that in this area there is a motor representation of hand and mouth which plays a fundamental role in the generation of goal directed motor acts such as grasping [Gentilucci et al. (1988); Kurata and Tanji (1986); Maranesi et al. (2012); Rizzolatti et al. (1988)]. Different subsections of area F5 (F5p and F5c) are strongly connected with the parietal areas AIP and PFG. Pani et al. (2014) showed that area AIP contains neurons (mirror) that are active during both observation and execution, suggesting that AIP-F5 connections could represent a component of the mirror system related to the description of hand-object interactions. Area PFG is mostly a hand related field which contains neurons coding goal directed motor acts as well as mirror neurons [Fogassi et al. (2005); Rozzi et al. (2008)]. In this area, the activity of a large portion of these neurons appears to be involved in higher-order aspects of action organization related to the motor intention of the agent [Rizzolatti and Fogassi (2014)]. Furthermore, fMRI was used to examine activation of superior temporal sulcus (STS) to understand how visual responses related to action observation might reach the premotor areas. It has been shown that observing actions activates the dorsal and ventral banks of the superior temporal sulcus (STS), projecting to the parietal motor areas PFG and AIP respectively [Nelissen et al. (2011)]. In the same study, another area revealing significant action observation fMRI responses was the prefrontal area 46. A possible role of the prefrontal cortex in the mirror system could be to provide motor representations of the parietal an motor areas with contextual information, for example allowing action understanding when the whole action is not visible [Umiltà et al. (2001)]. After the discovery of mirror neurons in monkeys, evidence that observing actions performed by others recruits a pattern of brain areas typically activated during the execution of those actions have also been demonstrated in humans (Figure 2.2) [Keysers and Gazzola (2009); Rizzolatti et al. (2014)]. As several meta-analyses have shown, the main nodes of human mirror system, in the case of grasping actions, are the inferior parietal lobule (IPL) and the ventral premotor cortex (PMv) plus the caudal part of the inferior frontal gyrus (IFG) [Caspers et al. (2010); Grosbras et al. (2012); Molenberghs et al. (2012)].

2.1.2 "WHAT": mirror neurons and goal coding

A large number of studies in both monkeys and humans showed that mirror neurons typically encode the action goal, i.e. the outcome to which an action is directed (for example grasping or passing an object) [Rizzolatti et al. (2014)]. Interestingly, studies in monkeys showed that the parieto-frontal mirror neurons respond to the goal of an action when the action is performed with different effectors [Ferrari et al. (2005); Gallese et al. (1996)] or tools [Rochat et al. (2010)]. Mirror neurons can generate the representation of a goal even when the monkey does not move but simply observes an agent moving to reach that action goal. Indeed, observing an action elicits a brain motor representation of the outcome to which the action is directed (its goal), allowing to identify the observed action goal relying on own motor processes [Rizzolatti et al. (2001); Rizzolatti and Sinigaglia (2010)]. Two studies demonstrated this point. In a first study, mirror neurons responses were recorded while the monkey observed grasping actions performed by an experimenter. Particularly, in one condition ("hidden condition") the final part of the grasping action (the one involving the hand-object interaction) occured behind an opaque screen and thus was not visible to the monkey [Umiltà et al. (2001)]. Results showed that the majority of mirror neurons responded also in the hidden condition. This response can be considered a "mental operation" which

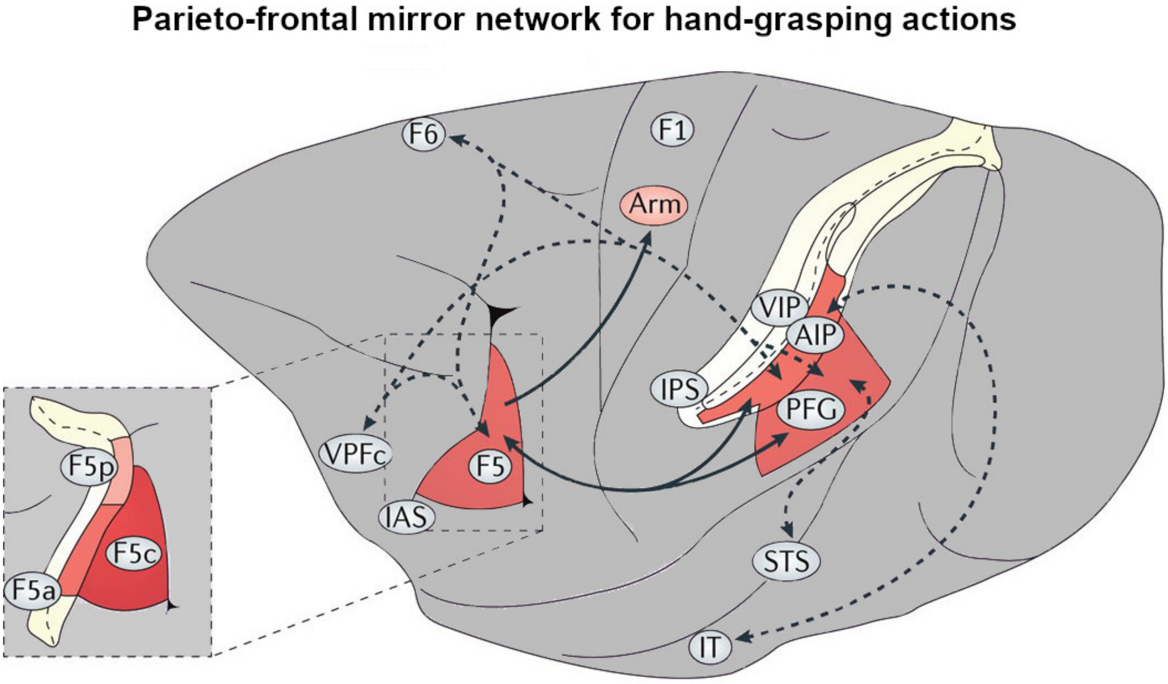


Figure 2.1 Monkey parietofrontal mirror network for hand-grasping actions involving: the ventral premotor cortex (area F5), area PFG and anterior intraparietal area (AIP). The parietofrontal network receives high-order visual information from areas of the temporal sulcus (STS) and inferior temporal lobule (IT) and is under the control of frontal lobe (F6 and VPFc). Area F5 is connected with area F1. Figure adapted from Rizzolatti and Sinigaglia (2016).

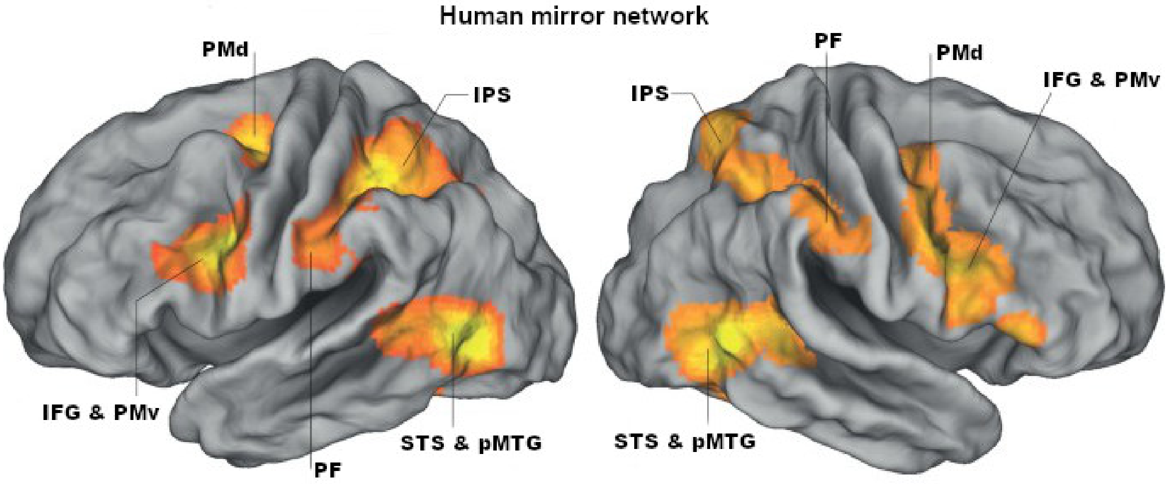


Figure 2.2 Human grasping network. Activations are displayed on the left and right lateral-surface view of the MNI single-subject template: inferior frontal gyrus (IFG), intraparietal sulcus (IPS), dorsal premotor cortex (PMd), posterior middle temporal gyrus (pMTG), ventral premotor cortex (PMv), parietal area F (PF), superior temporal sulcus (STS). Figure adapted from Rizzolatti and Sinigaglia (2016).

enabled to internally reconstruct the hidden motor act on the basis of its stored motor representation. In a second study, Kohler et al. (2002) showed that "audiovisual" mirror neurons respond not only during the observation/execution of an action (e.g. paper ripping) but also to the sound produced by that action. This suggests that the representation of a specific action goal can be accessed by multiple sensory modalities.

2.1.3 "WHY": mirror neurons and intention coding

Following the description of Berstein (1996), who defined actions as sequences of motor acts that all together aim to reach a final motor goal, we can identify at least two levels of goal coding: the immediate goal (e.g. grasping an apple) and the final goal (grasping an apple *to eat it*). The final goal specifies *why* the action is performed, its motor intention. Behavioural studies in humans showed that the final goal of the whole action influences the kinematic features of each motor act of the sequence and that the specific plan of each motor act changes depending on the act following it within the action sequence [Gentilucci et al. (1997); Marteniuk et al. (1987); Rosenbaum et al. (2012)]. Cohen and Rosenbaum (2004) showed that when individuals reached out to take hold of a plunger to move it to a target location, they grasped the plunger low for high target locations and high for low target locations. This "grasp height" effect is an example of how people take end-states into account in movement planning. Interestingly, in a subsequent experiment, Rosenbaum et al. (2006) showed that also the width of the target site influenced the observed grasp heights. Particularly, when the target ring was narrow, participants grasped the plunger lower than when the target ring was wide. These outcomes indicate that the movement as a whole, or at least aspects of its start and end, was known by the time the plunger was grasped for movement. The intention to achieve a given goal implies an organization of the entire chain of motor acts forming the action.

An interesting point is to understand if mirror neurons code the single motor act or their response might change according to the whole action in which that motor act is embedded. Fogassi et al. (2005) demonstrated this point by recording the activity of neurons in area PFG while monkeys performed two grasping actions: grasping a piece of food and eat it and grasping an object and place it into a container. Results showed that the discharge of PFG neurons not only encodes motor acts in terms of their immediate goal (grasping) but reflects the final goal of the whole action (eating vs. placing). The same results were confirmed for area F5, which is strongly connected with PFG [Bonini et al. (2010); Rozzi et al. (2006)]. Since the activity of neurons in both areas is modulated depending on why the monkey is grasping the object, it is possible to claim that their discharge reflects the monkey's intention. Bonini and colleagues demonstrated that, also for more complex sequences of motor acts, motor intention does not terminate with movement onset but guides action unfolding until the action completion [Bonini et al. (2011); Pacherie (2008)]. In the same studies described above, authors recorded the activity of grasping neurons not only during the execution but also while the monkey observed similar actions performed by the experimenter. Results showed that, also for observation, the discharge intensity of all recorded mirror neurons was modulated by the final goal of the observed action.

2.1.4 The role of motor expertise in action understanding

Several behavioural and neuroimaging studies showed the crucial role of *motor expertise* for the understanding of others' actions: the greater the motor expertise of individuals and the greater their ability to activate the mirror mechanism and identify the outcome to which those actions are directed. For example, an fMRI study comparing the brain activity when expert dancers watched their own dance style versus another style revealed greater bilateral activations in premotor cortex and intraparietal sulcus, right superior parietal lobe and left posterior superior temporal sulcus, suggesting that the mirror system integrates observed actions with an individual's personal motor repertoire [Calvo-Merino et al. (2005)]. It has been systematically demonstrated that the recruitment of motor areas with mirror properties strongly correlates with motor rather than visual expertise [Calvo-Merino et al. (2006); Cross et al. (2009); Kirsch and Cross (2015)]. Another example comes from a study which combined physchophysical and neurophysiological techniques (TMS) [Aglioti et al. (2008)]. Participants, including professional basketball players (visuomotor experts), coaches and journalists (visual experts) and novice individuals, were asked to observe videoclips showing free shots at a basket and judge if the shots would be successfull or not. Interestingly, the experiment was characterized by a temporal occlusion paradigm: videoclips could stop after about 400-500 ms when players had the ball still in their hand, after 700-800 ms when the ball began its upward trajectory towards the hoop, or after about 1200 ms when the ball started its downward path towards the hoop. Results showed that former players (visuomotor experts) were quicker and more accurate in the task compared to both visual experts and novices, and showed a time/muscle specific modulation of covert motor activation during the observation of erroneous throws. These data indicate that motor expertise may weigh more than visual expertise alone in predicting the outcome of the observed action.

Taken together, findings reviewed so far identify neurophysiological bases and functional properties of the mirror mechanism, suggesting that, far from representing a peculiar feature of premotor cortex only, it represents a basic principle of brain functioning, enabling to *understand others from inside.* Observing a person who performs an action allows us to understand *what* this person is doing (e.g. grasping a cup) and *why* (e.g. grasping a cup for drinking or passing it to another person). However, as suggested by the American psychologist Daniel Stern, a third fundamental component has been neglected from neuroscience studies and particularly from mirror neurons research framework, despite its pivotal role in social communication: *how* the action is performed. The following section will leverage on Stern's words and research to introduce the main topic of my PhD: *vitality forms*.

2.2 What are Vitality forms?

The American psychoanalyst and developmental psychologist Daniel Stern defined vitality forms starting from the long-term observation of the affective bonding relations of infants with their mothers [Stern (1977, 1985)]. In his book "The Interpersonal World of the Infant", Stern gives an example describing a child of ten months sitting on the floor facing her mother while trying to fit a piece of a puzzle in its right place:

"She looks up into her mother's face with delight and an explosion of enthusiasm. She 'opens up her face' [her mouth opens, her eyes widen, her eyebrows rise] and then closes back down. The time contour of these changes can be described as a smooth arch [a crescendo, high point, decrescendo]. At the same time her arms rise and falls at her sides. Mother responds by intoning, 'Yeah' with a pitch line that rises and falls as the volume crescendos and decrescendos: 'yeeAAAaahh.' The mother's prosodic contour matches the child's facial-kinetic contour. They also have the exact same duration".

Through this "affective attunement", the mother is able to understand the mood of her daughter and simultaneously maintain a correspondence of vitality in her response, even if conveyed through a different sensory modality. Thus, what are vitality forms?
Vitality forms are related to basic kinematic characteristics of actions, an external representation of our internal states. Specifically, besides action goal and intention, they represent a third fundamental action component: the "how" dimension. Depending on affective or cognitive states of individuals, their actions may take a specific kinematic "contour". For example, the action of grasping a cup can be "rude" or "gentle", an handshake can be

"vigorous" or "delicate" etc. As stressed by Stern, the concept of vitality forms refers to a spontaneous integration of different dynamic components (movement, force, space, time, direction/intention) linked and perceived together as a whole, a Gestalt [Stern (2010)].
Thus, a variation in one or more of these components gives rise to specific vitality forms that observers can capture from actions performed by others and through which they can infer their positive or negative attitudes.
Thanks to the collaboration of Daniel Stern, researchers at the Department of Neuroscience of Parma started the investigation of neural bases of vitality forms by using fMRI as main neuroimaging technique. A challenging issue consisted of understanding how to dissociate the two components of actions observed/executed by participants (goal and form) in order to identify only brain regions selective for vitality forms processing. As suggested by Stern, one way consisted in varying the goal (e.g. taking vs giving) keeping the same form or varying the vitality form (gentle vs rude) keeping the action goal constant. Moreover, given the significance of vitality forms in social interactions, Stern suggested to consider an interpersonal context where actions were not merely directed towards objects but towards other individuals.

2.3 Vitality forms: from neuroscience to robotics

2.3.1 Neural bases: fMRI experiments

The first evidence for the localization of neural correlates specific for vitality forms was obtained in a fMRI study performed by Parma team in collaboration with Daniel Stern [Di Cesare et al. (2013)]. In this study participants were presented with video-clips showing two actors, sitting at a table facing each other, performing actions (Figure 2.3 A-B) with two different vitality forms (gentle or rude). In particular, four transitive actions (grasping a cup, passing a bottle, giving a packet of crackers, passing a ball) and four intransitive actions (clapping hands, shaking hands, stroking the other actor's backhand and stopping gestures) were presented in pairs of consecutive video-clips in which the action goal (What) and the action form (How) could be the same or could change between the video pairs. Participants performed two tasks: in the "What" task, they were asked to pay attention to the action goal and decide whether the two actions were the same or different, regardless of their vitality form; in the "How" task, they were asked to focus on the action form and decide whether the vitality forms were the same or different in the two consecutive videos, regardless of the type of action performed. Activations resulting from both tasks were found in the parieto-frontal circuit, typically involved in the observation and execution of actions (Figure 2.3 C) The most relevant result derived from the contrast "How" vs "What", which showed a selective activation of the dorso-central insula (DCI) (Figure 2.3 D).

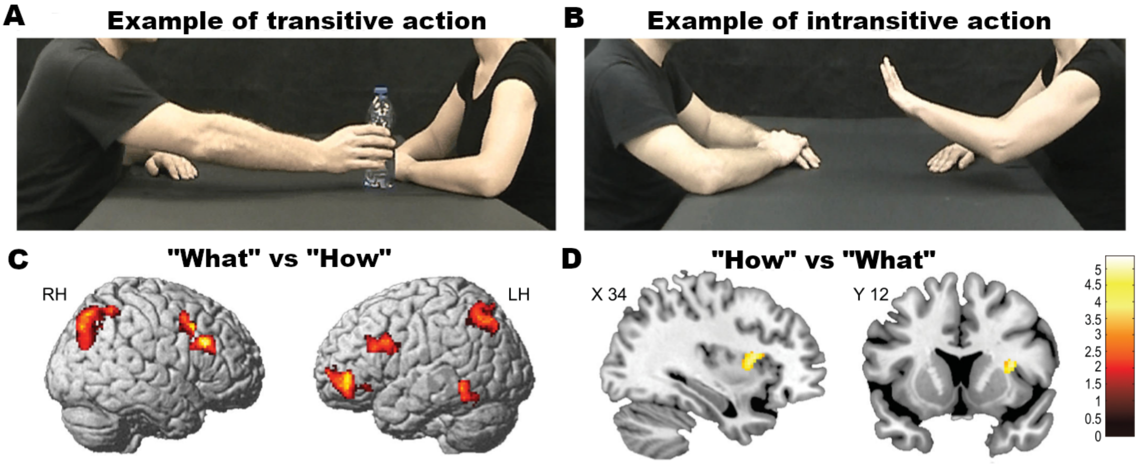


Figure 2.3 Example of transitive action (passing action) (A). Example of intrasitive action (stop gesture) (B). Activations resulting from the contrast "What vs How" (C) and "How" vs "What" (dorso central insula, D). Figure adpted from Di Cesare et al. (2013)

A question arised immediately after: what happens when an action conveying vitality forms is imagined or executed in first person and not just observed?

To address this issue, in a subsequent fMRI study participants were required to perform three different tasks: observation (OBS), imagination (IMA) and execution (EXE) [Di Cesare et al. (2015)]. In the OBS task, they observed video clips showing the right arm of an actor performing actions toward another actor, with a gentle or a rude vitality form. In the IMA task, participants had to imagine themselves performing the actions seen during the observation task, again with a gentle or rude vitality form. In the EXE task, participants moved a packet of crackers located on a plane lying on their chest as if offering them to another person, with a gentle or rude vitality form. For the same tasks, control conditions were created: observe actors positioning a small ball inside a box, imagine themselves performing the same action or execute the action inside the scanner. The main result of this study was the demonstration that, besides the observation, also the execution and the imagination of actions performed with a vitality form, relative to control actions, determines a specific activation of the dorso-central insula.

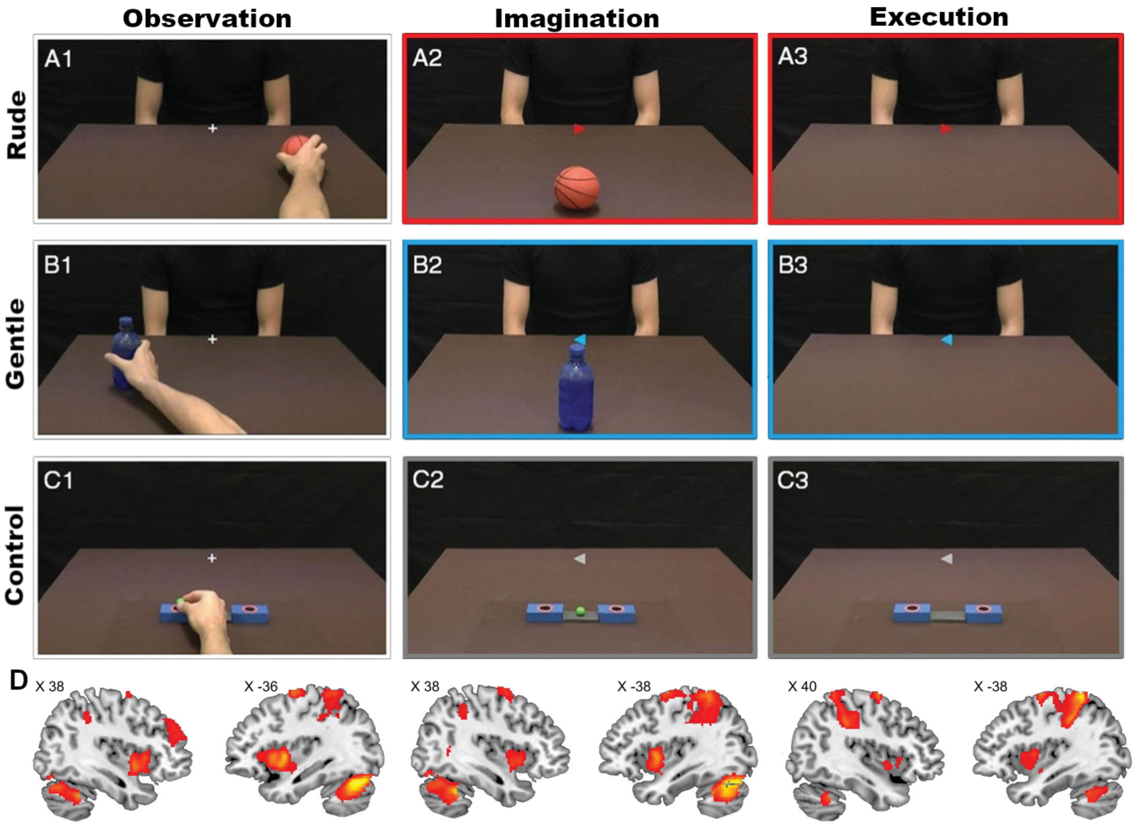


Figure 2.4 Experimental task design. Left panel, observation task for vitality forms condition (rude, A1; gentle, B1) or control condition (C1). Middle panel, imagination task for vitality forms condition (rude, A2; gentle, B2) or control condition (C2). Right panel, execution task for vitality forms condition (rude, A3; gentle, B3) or control condition (C3). Parasagittal sections showing the insular activations in the three conditions (D). Figure adapted from Di Cesare et al. (2015).

In a subsequent fMRI study, Di Cesare et al. (2019a) investigated whether listening to action sounds conveying vitality forms activates the same insular sector. The idea underlying this experiment was that every action performed with a certain vitality form produces a characteristic sound that tells us if the attitude of the agent is positive or negative. Participants were presented with five sounds coming from daily actions (e.g. stirring coffee, closing a door etc., Figure 2.5 A). In a condition, actions which produced sounds were characterized by rude or gentle vitality forms, while in another condition the vitality form conveyed by the action sound was masked (control condition). In agreement with previous data, listening to action sounds produced a bilateral activation of the parieto-frontal circuit known to be involved in action execution [Gazzola et al. (2006)]. Listening to the control stimuli also produced the activation of the same circuit indicating that, although information concerning vitality forms was absent, participants understood the action goals. The main result was that listening to gentle and rude action sounds produced the activation of the left dorso-central insula. Moreover, results indicated a selective activation of the middle cingulate cortex (MCC) during the vitality forms condition (Figure 2.5 B). Whereas the activation of DCI confirmed previous findings, the activation of MCC was unexpected.

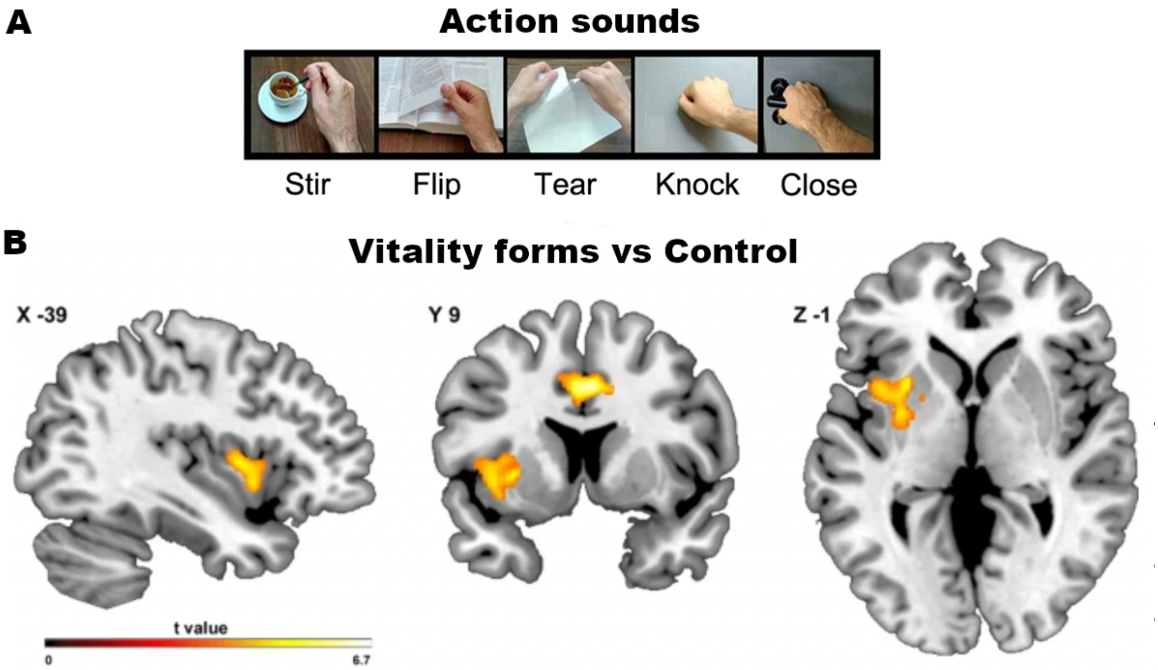


Figure 2.5 Actions sounds presented to participants (A). Vitality forms vs Control contrast revealing selective activation of DCI and MCC (B). Figure adapted from Di Cesare et al. (2019a).

2.3.2 Behavioural responses: a kinematic experiment

When an agent performs actions or pronounces words gently or rudely, a receiver immediately understands whether that agent is in a positive or negative mood. It is intuitive, therefore, that during social interactions vitality forms expressed by an agent may accordingly influence the behaviour of a receiver. A first demonstration of this effect was recently provided by a kinematic study [Di Cesare et al. (2017)]. During the experiment participants were presented with videoclips showing an actor or an actress performing two types of requests: a *giving* request (asking for a bottle, Figure 2.6 A) or a *taking* request (handing a bottle, Figure 2.6 B). Particularly, each request could be performed in a gentle or rude way in three different modalities: visual (V), auditory (A) or mixed modality (AV). After each request, participants had to perform a reaching-grasping action with the intention to give the bottle (if the request was giving) or take it (if the request was taking).

Results showed that, for both tasks and independently from the modality of expression, the perception of vitality forms characterizing the request modulated the kinematic features of participants motor response. Particularly, during the reaching phase vitality forms influenced spatial and temporal parameters showing a wider trajectory and higher velocity in response to rude requests compared to the gentle ones. Additionally, during the grasping phase, results indicated a wider hand aperture in response to rude vitality forms. These data clearly indicate that vitality forms expressed by others significantly modulate our consequent motor behaviour.

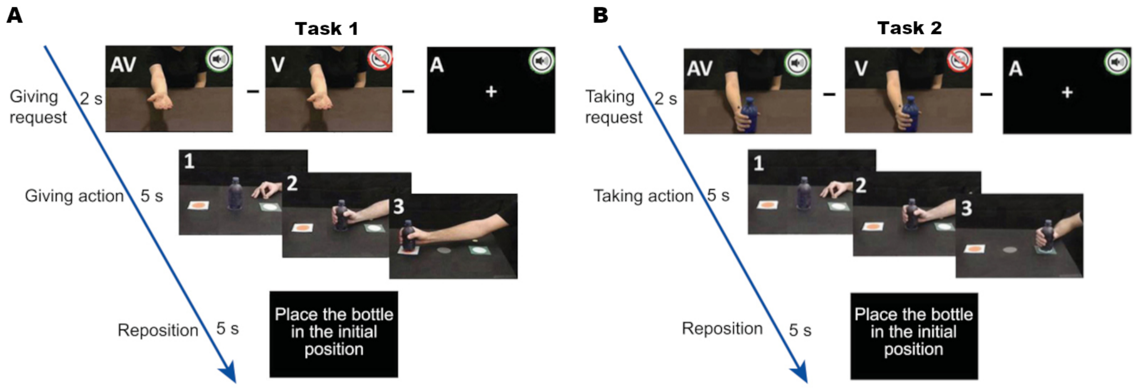


Figure 2.6 Experimental paradigm. Participants were presented with audio-visual (AV), visual (V) and auditory (A) stimuli. In task 1, participants were requested to give the bottle after a giving request (A). In task 2, participants were requested to take the bottle after a taking request (B). The timeline reports the timing of different trial phases. Figure adapted from Di Cesare et al. (2017).

2.3.3 Human—robot interactions

Research in robotics has often focused on generation of human-like movements in the attempt of creating communicative actions [Gielniak et al. (2013); Karg et al. (2013); Sandini et al. (2017); Sciutti et al. (2018)]. For example, different authors proposed to generate humanoid motions on the basis of the Laban Movement Analysis that describes the emotional component of movement by using parameters such as velocity, curvature and acceleration [Takayama et al. (2011)]. Considering the numerous findings in neuroscience, in recent years robotics started to understand the potential role of vitality forms in human—robot interactions. Indeed, the opportunity of humanoid behaviour to express vitality forms enriches the array of nonverbal communication that can be exploited by robots to foster an effective communication with humans. In this regard, the main challenges addressed consisted of 1) endowing a humanoid robot (the iCub robot) with vitality forms allowing it to generate gentle and rude actions; 2) investigating whether and how the observation of these actions influence the perception and motor behaviour of the human partner. Particularly, may the observation of actions generated by the iCub robot with a gentle or rude vitality form elicit the activation of the dorso-central insula and influence action execution as found during the observation of human actions?

To generate vitality forms on the iCub robot, a methodology which has been proven to be effective is the retargeting of human kinematic features onto the motor space of the robot. Particularly, a female actress whose arm size was similar to the one of the iCub robot was asked to perform an offering gesture gently or rudely. The kinematic parameters (i.e. peak velocity, length trajectory) of her actions were recorded with a motion capture system and remapped into the joint space of the iCub robot. This allowed to obtain an exact replica of human actions with the same vitality forms (Figure 2.7). Once endowed the iCub robot with the capability to express gentle and rude vitality forms with its arm, the next step consisted in validating these actions by investigating the effect on human participants.

Firstly, an fMRI study showed that the observation of robotic actions endowed with human vitality forms produced a BOLD signal increase in the dorso-central insula, very similar to that obtained during the observation of human actions. Indeed, the ROI analysis carried out in the dorso-central insula did not reveal any significant difference of the BOLD signal between the observation of human and robotic actions in both gentle and rude conditions (Figure 2.8 C). This result was not obtained in a previous fMRI study in which the observation of robotic actions performed with a "biological motion" signature did not elicit insular activity [Di Cesare et al. (2020b)]. In other words, modulating the velocity profile of robot actions according to the 2/3 power law does not produce vitality forms. Remarkably, kinematic features of human actions play a fundamental role in triggering the activation of the dorso-central insula.

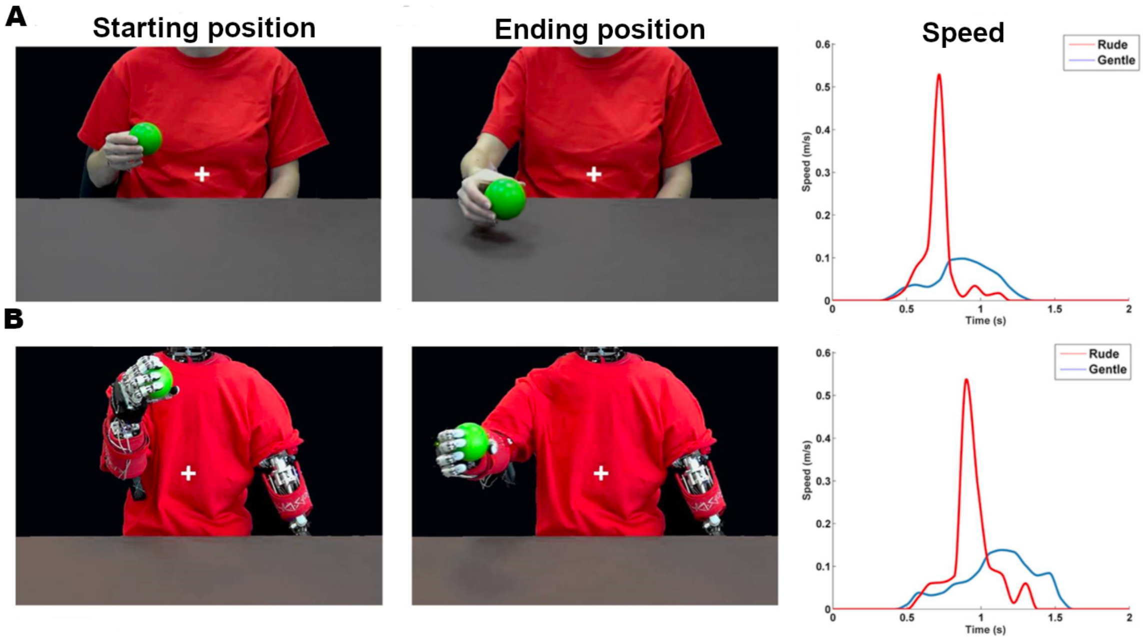


Figure 2.7 Starting position, ending position and speed characterizing actions of the human actress (A) and the iCub robot (B). Figure adapted from Di Cesare et al. (2020b).

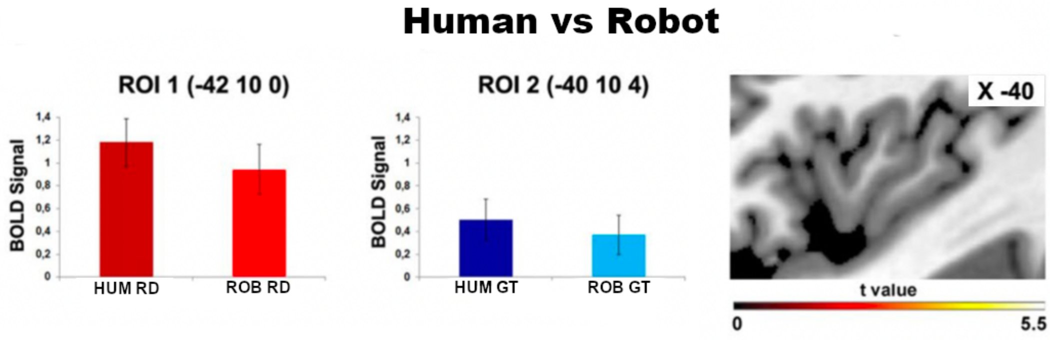


Figure 2.8 No significant difference of the BOLD signal during the observation of rude (left panel) and gentle actions (middle panel) perform by the human actress and the iCub robot. Contrast Human vs Robot in the dorso central insula (right panel). Figure adapted from Di Cesare et al. (2020b).

Secondly, a kinematic study assessed if the generated movements on the iCub robot effectively convey rude and gentle vitality forms to participants and consequently modify their motor behaviour. The iCub robot performed a pointing action in a gentle/rude way (giving request). After this request, participants reached the pointed object and passed it towards the iCub robot. To assess the effect of vitality forms on participants' motor response, kinematic features of their movement were recorded and analyzed. Results showed that participants were influenced by the vitality form expressed by the iCub robot. Particularly, their movement features resembled the ones of a rude action after the observation of a rude motion of the robot, while on the contrary they performed more gentle action when the robot was gentle itself.

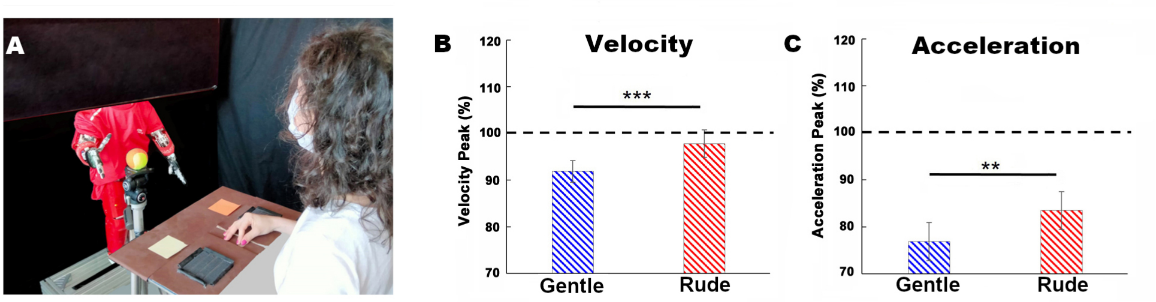


Figure 2.9 Experimental setup (A). significant effect of vitality forms expressed by the iCub robot on velocity (B) and acceleration (C) of participants motor response.

The present chapter reviewed the literature and research background of my thesis, starting from an in-depth analysis of the mirror mechanism and its role in human actions understanding. It then shifted focus towards the presentation of vitality forms as a fundamental action component for human—human interactions. Firstly, it delved into the earliest findings on neural correlates of vitality forms in humans, progressing to first insights regarding their behavioural consequences and finally exploring their potential application in human—robot interactions. Next chapters will present the experiments carried out during my PhD. Following a framework consistent with the state of the art here presented, we will first go through a deeper exploration of brain mechanisms underlying the processing of vitality forms in humans and secondly we will provide a more detailed understanding of the influence of vitality forms in both human—human and human—robot interactions.

Part I

Neural bases

Chapter 3

A mirror circuit encoding observation and execution of vitality forms

The outcomes presented in the following chapter have been all published in Proceedings of the National Academy of Sciences (Di Cesare et al. (2021a)). In this study, I personally contributed to performing the experiment, analysing data, interpreting-discussing results and writing the final paper.

3.1 Introduction

Despite the crucial role of vitality forms in social interactions, very little is known about their neurophysiological bases and only in recent years have data started to be collected on this issue. The previous Chapter reviewed the main results of the first fMRI studies aiming to assess the neural correlates underlying the processing of actions conveying vitality forms. Starting from these data, some insightful points still need to be addressed.

Firstly, results showed that the dorso-central sector of the insula (DCI) is the region selectively active during the observation and execution of actions with vitality forms. The overlap of perception and execution suggests that DCI is endowed with mirror properties. However, there are no results that computationally demonstrate in one experiment that the same voxels of DCI involved in the observation of action vitality forms are also involved during their execution, providing a clear evidence of the presence of a mirror mechanism specific for vitality forms processing.

Secondly, as described in the previous Chapter, it has been recently shown that also the middle cingulate cortex (MCC) might be involved in the encoding of vitality forms. Specifically, it was found activated during the listening of action sounds conveying rude and gentle vitality forms. While the activation of DCI has been extensively described from several studies, the selective activation of MCC was unexpected and no further studies have ascertained its possible role during vitality forms observation and execution.

In this view, the fMRI study presented in this Chapter aims to deepen the investigation of neural bases of vitality forms, by finding out the role of MCC, in addition to DCI, and demonstrating the existence of a mirror mechanism specific for vitality forms in both these brain areas.

To this purpose, participants were required to perform two tasks:

1) to observe an arm action (*Observation task*) performed gently and rudely (vitality forms condition) or with no vitality forms (using jerky movements as a control condition)

2) execute the same action (*Execution task*) gently and rudely (vitality forms condition) or with no vitality forms (control condition).

Once having identified brain areas involved in both tasks, a voxel based analysis was carried out to quantify the amount of voxels similarly activated during the observation and execution of actions conveying vitality forms. Finally, a multifiber tractography analysis was used to identify anatomical connections between active sites.

3.2 Methods

3.2.1 Participants

Sixteen healthy right-handed volunteers took part in the main fMRI experiment (nine females and seven males, mean age = 25.4, SD = 2). Due to the COVID-19 pandemic, only fourteen participants of the same group took part in the second scanning session for the collection of diffusion tensor imaging (DTI) images. All participants had normal or corrected-to-normal visual acuity. None of them reported a history of psychiatric or neurological disorders or current use of any psychoactive medications. All of them gave their written informed consent to be subjected to the experimental procedure, which was approved by the Local Ethics Committee of Parma (552/2020/SPER/UNIPR) in accordance with the Declaration of Helsinki.

3.2.2 Experimental paradigm

The fMRI experiment was based on a 2 x 2 factorial design with task (observation, OBS; execution, EXE) and conditions (Vitality forms, VF; Control, CT) as factors of interest.

In the OBS task, participants were presented with video-clips showing the right hand of an actor performing leftward passing actions of four different objects (a packet of crackers, a ball, a cup, a bottle) towards another actor sitting in front of the first one. These passing actions were performed by using two different vitality forms (gentle and rude, VF OBS) or in a jerky way (control condition, CT OBS). Specifically, the jerky actions used as control stimuli were obtained by presenting one static frame of the action every 700 ms (4 frames in total from the beginning to the end of the action). Thus, the control stimuli allowed participants to understand the action goal (passing an object) without conveying any vitality form information.

In the EXE task, participants were presented with the static image of an actor sitting in front of them and were required to move a little object (a little box with dimensions 7 cm x 3.5 cm located on a plane lying on their chest inside the scanner) with the intention to offer it to the actor with vitality forms (gentle or rude, VF EXE) or in a neutral way without vitality forms, by simply rotating the wrist (control condition, CT EXE). A cue positioned in the center of the screen indicated when to start the action. The colour of the edge of the screen indicated the way in which participants were required to perform the action (blue: gentle; red: rude; gray: control). In each video, a fixation cross was introduced to control for restrained eye movements. Video clips were recorded in a dark scenario and actors wore black shirts to emphasize the forelimbs. Additionally, in order to focus the attention of participants on the performed action and exclude the vision of the actor's face, original videoclips were cropped by using VirtualDubMod software v1.5.

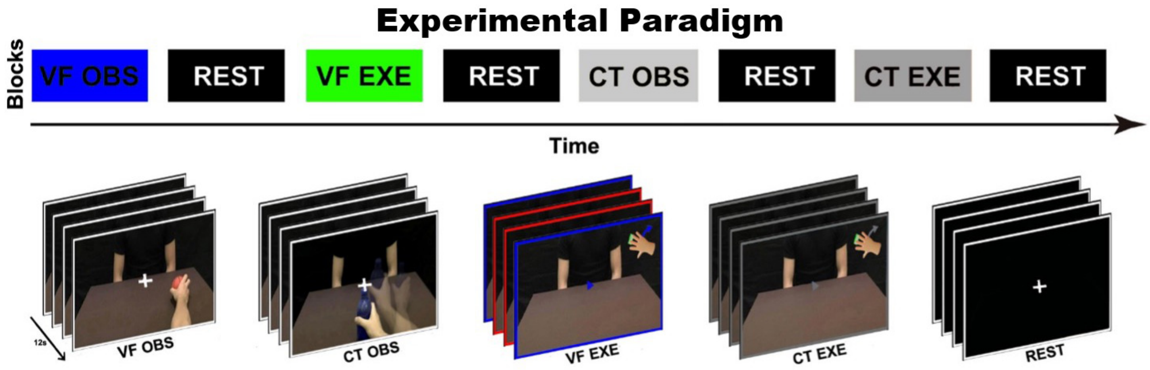


Figure 3.1 In the Observation task, participants were asked to observe actions conveying vitality forms (VF OBS) or jerky actions (CT OBS). In the Execution task, participants were asked to perform the action with (gently: blue colour, rudely: red colour; VF EXE) or without (neutrally: grey colour; CT EXE) vitality forms. Between blocks, a rest period of 12 s was presented (REST).

The experiment was presented as a blocked design and subdivided in 2 runs. Particularly, in each run, stimuli were presented in mini-blocks of four consecutive videoclips (each lasting 3 s) of the same condition (conditions: VF OBS, CT OBS, VF EXE, CT EXE). Each run presented 4 blocks for condition in a randomized order. During the OBS task, participants observed 32 actions in total for each condition (VF OBS, CT OBS), while for the EXE task they performed 32 actions for each condition (VF EXE, CT EXE). Moreover, the vitality form (gentle, rude) of the observed and performed actions was randomly changed in order to avoid a possible repetition-suppression effect. Between experimental blocks, participants were asked to observe a white fixation cross on a black screen (rest period lasting 12 s). Stimuli were presented via digital visors (VisuaSTIM) with a 500,000 px x 0.25 square inch resolution and horizontal eye field of 30°. The digital transmission of the signal to the scanner was via optic fiber. The software E-Prime 2 Professional was used for stimuli presentation.

3.2.3 fMRI Data Acquisition

Participants lay in the scanner in a dimly lit environment. Anatomical T1-weighted and functional T2*-weighted MR images were acquired with a 3 Tesla General Electrics scanner equipped with an 8-channel receiver head-coil. Functional images were acquired by using a T2*-weighted gradient-echo, echo-planar (EPI) pulse sequence acceleration factor asset 2, 40 sequential transverse slices (slice thickness = 3 plus inter-slice gap = 0.5 mm) covering the whole brain, with a TR time of 3 s. Each functional run comprised 169 ascending interleaved volumes. Additionally, a T1 weighted structural image was acquired for each participant.

3.2.4 fMRI Data Analysis

Analysis of fMRI data was performed with SPM software running on MATLAB (R2018b) [Penny et al. (2007b)]. After standard preprocessing steps, data were analyzed using a random-effects model, implemented in a two-level procedure. In the first level, single-subject fMRI BOLD signal was modelled in a General Linear Model (GLM) comprising the following regressors: Vitality Forms Observation (VF OBS), Control Observation (CT OBS), Vitality Forms Execution (VF EXE), Control Execution (CT EXE) and Instruction. Within each block, video-clips were modelled as a whole event lasting 12 s. The instruction was modelled with a duration of 3 s.

In the second level analysis (group-level analysis), corresponding contrast images of the first level for each participant were entered into a flexible ANOVA with sphericity-correction for repeated measures. This model was composed of 4 regressors: VF OBS, CT OBS, VF

EXE, CT EXE. Within this model, we assessed activations associated with each task versus implicit baseline and activations resulting from the direct contrast between conditions (VF OBS vs. CT OBS, VF EXE vs. CT EXE; pFWE <0.05 corrected at the cluster level).

In order to evaluate and compare the BOLD signal during the processing of gentle and rude vitality forms, we carried out an additional analysis. Specifically, at first level we modeled the BOLD signal change by using another GLM which comprised the following regressors: Gentle Vitality Forms Observation (GT OBS), Rude Vitality Forms Observation (RD OBS), Control Observation (CT OBS), Gentle Vitality Forms Execution (GT EXE),), Rude Vitality Forms Execution (RD EXE), Control Execution (CT EXE) and Instruction. Within each block, video-clips were modelled as single event lasting 3 s as well as the instruction.

In the second level analysis (group-level analysis), this model was composed of 6 regressors: GT OBS, RD OBS, CT OBS, GT EXE, RD EXE, CT EXE. Within this model, we assessed activations associated with each task and condition versus implicit baseline (pFWE <0.05 corrected at the cluster level).

A conjunction analysis of the brain activations resulting from the contrasts VF OBS vs. CT OBS and VF EXE vs. CT EXE (pFWE < 0.05 corrected at the cluster level) was carried out to highlight brain areas selectively active during both observation and execution of actions conveying vitality forms. For each participant and for each single voxel, the BOLD signal change relative to each condition (VF OBS, CT OBS, VF EXE, CT EXE) was extracted in regions of interest resulting from the conjuction analysis by using the REX toolbox [Duff et al. (2007)]. Subsequently, the BOLD signal change (average value) obtained in the VF OBS condition was correlated with that obtained in the VF EXE condition in order to identify and quantify voxels showing a very similar (strong correlation) BOLD signal change during the observation and execution of actions conveying vitality forms, i.e voxels endowed with mirror properties and characterized by high correlated activity (HC mirror voxels). Only voxels showing at least 50% of the explained variance were considered. Additionally, a cutoff correlation value at $r > 0.7$ was used (coefficient of determination $R^2 \geq 0.49$, $p < 0.05$).

3.2.5 DTI Data Acquisition and Analysis

In another scanning session, diffusion tensor imaging data (DTI) were acquired on the same group of participants. Specifically, a diffusion spin-echo single shot echo planar imaging sequence with 64 diffusion directions (effective b-value of 1,000 s/mm2), 8 images with no diffusion weight in the anterior–posterior phase encoding direction and 8 images with no diffusion weight in the reverse phase encoding direction were collected. FMRIB Software

Library (FSL) tools (version 5.0.9) was used for DTI data processing. The first step (TopUp) of the preprocessing phase consisted in estimating the effective magnetic field voxel per voxel, from the volumes acquired in two different modalities (Anterior-Posterior and Posterior-Anterior). Particularly, by using acquisitions with opposing polarities, meaning that the same field leads to distortion going in opposing directions, TopUp estimates the field that, when applied to the two volumes, maximizes their similarity. This field was used in the second step (Eddy) for the correction of distortions caused by induced currents and any movements occurring during the acquisition. The third step (Bet Brain Extraction) resulted in a final image of the brain of each participant without the skull. A fundamental point of this third passage consisted in correctly adjusting the threshold (generally set to 0.5) to obtain a brain mask with a good coverage. The fourth step (Bedpost Analysis), preliminary to tractography, created the tensor of diffusion. Finally, in the fifth and last step, a probabilistic tractography was carried out with the FSL PROBTRACKX tool, which extracted the most likely tract between the two regions of interest (ROIs). The two ROIs were created as spheres with diameter of 12 mm and centered in coordinates characterized by good activation in fMRI group-level analysis during OBS and EXE tasks and located in white matter (DCI: x 32 y 9 z -2; MCC: x -9 y 9 z 42).

3.3 Results

The contrast VF OBS vs. CT OBS showed that the vitality forms condition produced a consistent activation of the left MCC with an extension to the pre-supplementary area (preSMA) bilaterally (pFWE - corrected 0.0001, Ke = 1477 voxels), the left DCI (pFWE - corrected 0.0001, Ke = 1486 voxels), the middle temporal gyrus (left hemisphere: pFWE - corrected 0.0001, Ke = 631 voxels; right hemisphere: pFWE - corrected 0.0001, Ke = 993 voxels) and the right premotor cortex extending to the inferior frontal gyrus (pFWE - corrected 0.0001, Ke = 2909 voxels). Similarly, the contrast VF EXE vs. CT EXE showed that the vitality forms condition produced the activation of the same sectors of the left cingulate cortex extending to the ride side (pFWE - corrected 0.0001, Ke = 1942 voxels) and the insula bilaterally (left hemisphere: pFWE - corrected 0.0001, Ke = 520 voxels; right hemisphere: pFWE - corrected 0.002, Ke = 328 voxels). Results of the conjunction analysis VF OBS vs. CT OBS and VF EXE vs. CT EXE revealed the activation of the left DCI (pFWE - corrected 0.001, Ke = 349 voxels) and the MCC with an extension to the preSMA (pFWE - corrected 0.0001, Ke = 826 voxels). Results of the correlation analysis used to highlight voxels showing a discreet/strong correlation between the two tasks are shown in

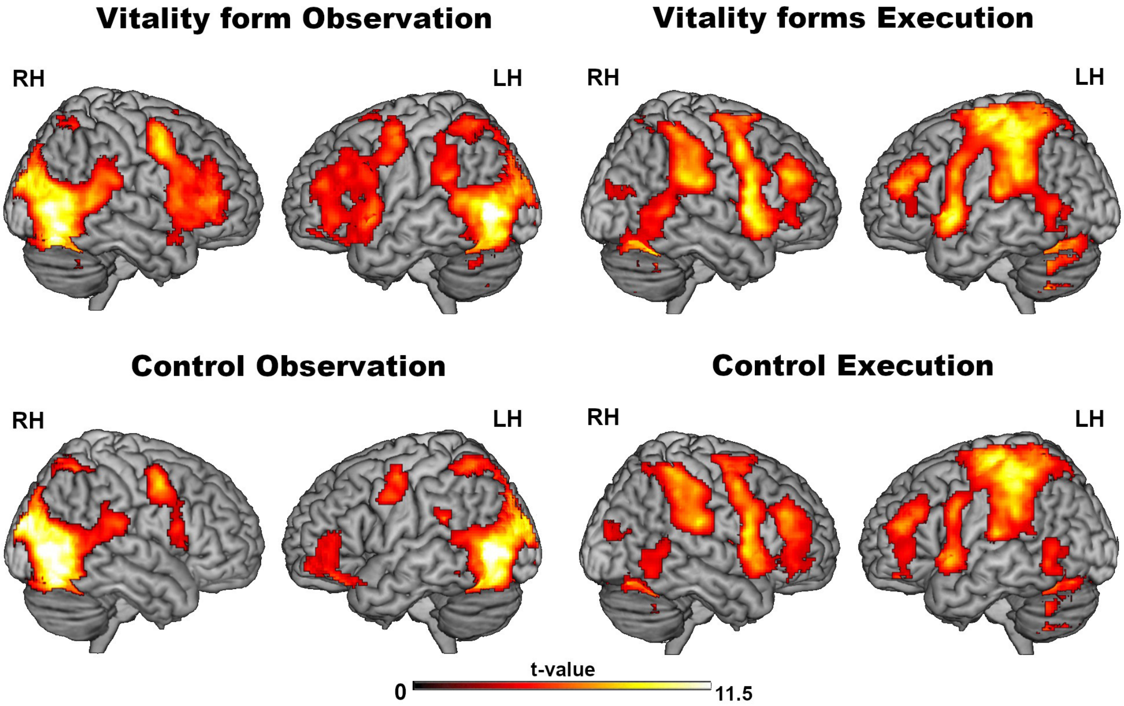


Figure 3.2 Brain activations resulting from vitality forms (VF) and control (CT) conditions vs. implicit baseline during the OBS and EXE tasks, rendered into a standard MNI brain template (PFWE<0.05 at voxel level). Figure adapted from Di Cesare et al. (2021a).

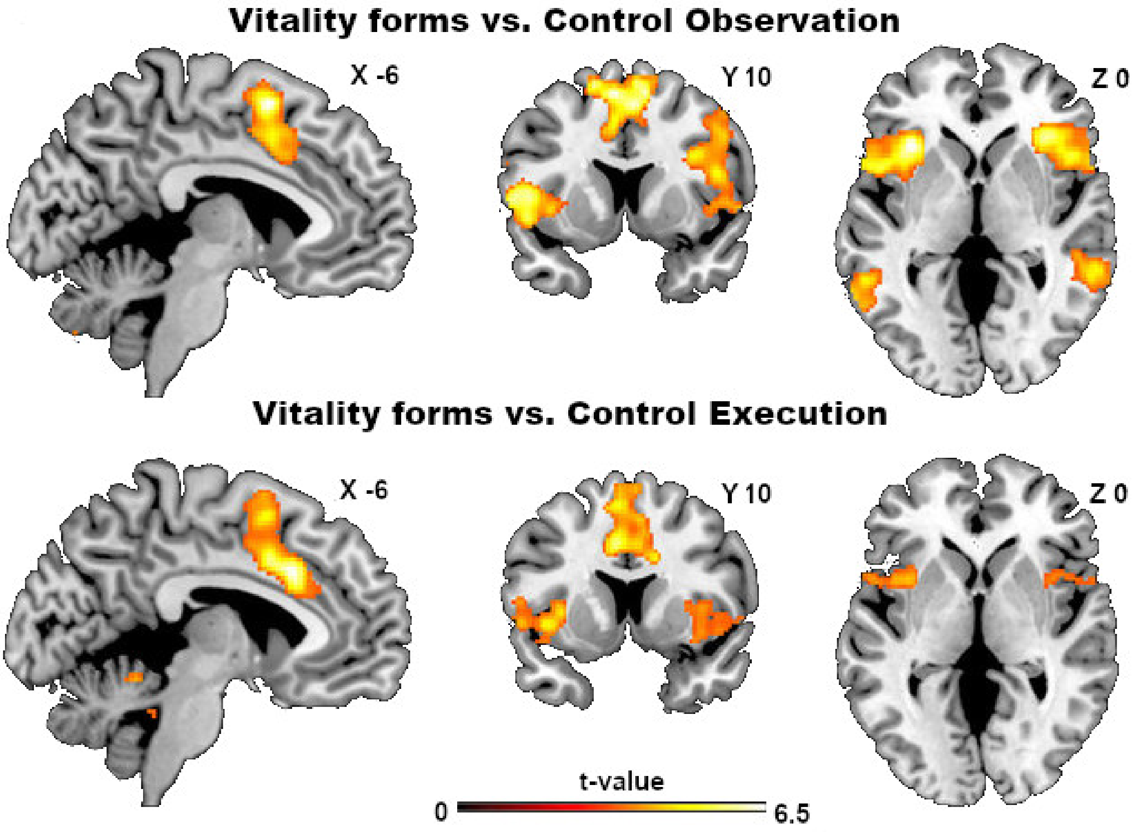


Figure 3.3 Sagittal and coronal sections showing the activation of the insular and cingulate cortices in the two hemispheres resulting from the direct contrasts VF OBS vs. CT OBS (first row) and VF EXE vs. CT EXE (second row). These activations are rendered using a standard Montreal Neurological Institute (MNI) brain template (PFWE < 0.05 at cluster level). Figure adapted from Di Cesare et al. (2021a).

Figure 3.5 A2-B2. Each square of the grids represents a single voxel. Orange squares indicate voxels selective for the encoding of vitality forms during both OBS and EXE tasks.

For the whole cluster of the insula (Figure 3.5 A1), results showed 147 out of 349 voxels (42.1%) with a significant correlation between vitality form task (VF OBS, VF EXE), 17 out of 349 voxels (4.8%) with a significant correlation between control tasks (CT OBS, CT EXE), whereas the remaining voxels (53.1%) with a weak correlation between tasks ($r < 0.7$). A further analysis was restricted to voxels located in the left DCI by applying an inclusive mask (Figure 3.5 A3) obtained from previous fMRI studies [Di Cesare et al. (2015, 2013, 2020a)]. Results of this analysis revealed 55 out of 140 voxels (39.2%) with a strong significant correlation between vitality form observation and execution (HC mirror voxels; $r > 0.7$, $p < 0.05$) and no voxels showed a significant correlation between control tasks.

For the whole cluster cingulum-preSMA (Figure 3.5 B1), results showed 397 out of 826 voxels (48%) with a significant correlation between vitality form tasks (VF OBS, VF EXE), 112 out of 826 voxels (13.5%) with a significant correlation between control tasks (CT OBS, CT EXE), whereas the remaining voxels (38.5%) with a weak correlation between tasks ($r < 0.7$). A further analysis was restricted to voxels located in the left MCC by applying an inclusive mask (Figure 3.5 B3) obtained from a previous fMRI study [Di Cesare et al. (2019a)]. Results of this analysis revealed 88 out of 181 voxels (48.6%) in the left MCC with a strong significant correlation between vitality form observation and execution (HC mirror voxels; $r > 0.7$, $P < 0.05$) and no voxels showed a significant correlation between control tasks. Furthermore, the comparison between gentle and rude conditions revealed a significant difference of the BOLD signal change during the OBS and EXE tasks in HC mirror voxels of both the MCC and DCI (paired sample t-test, $p \leq 0.05$, Figure 3.5 A4-B4).

Results of the tractography analysis revealed that DCI and MCC in the left hemisphere are anatomically connected. Particularly, a reconstruction of the average tract, obtained using a single tract from each subject (with 10% threshold), is shown on 2D and 3D brain templates (Figure 3.4 B and C).

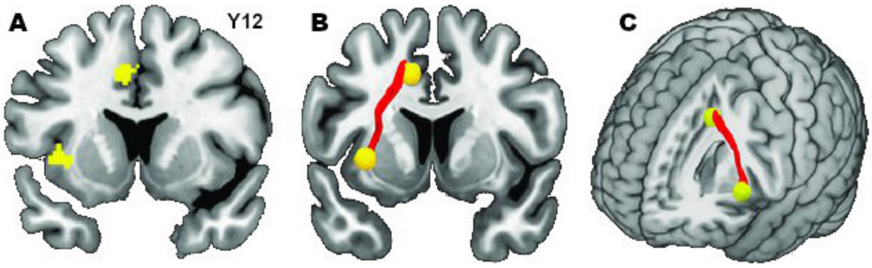


Figure 3.4 Anatomical connectivity between DCI and MCC. Activation of the left DCI and MCC resulting from the conjunction analysis of VF OBS vs. CT OBS and VF EXE vs. CT EXE contrasts (A). White-matter tract connecting the MCC and DCI in 2D view (B) and 3D view (C). Figure adapted from Di Cesare et al. (2021a).

3.4 Discussion

The main finding of the present study is that both the dorso-central insula (DCI) and middle cingulate cortex (MCC) are endowed with a mirror mechanism specific for vitality forms processing. Indeed, a voxel based analysis revealed a significant correlation of the BOLD signal during observation and execution tasks in many voxels located in the DCI and the MCC, i.e. HC (high correlation) voxels. Within the HC voxels of both areas, the BOLD signal change showed a stronger activity for rude vitality form compared to the gentle one. Finally, a multifiber tractography analysis showed that DCI and MCC are anatomically connected, forming a circuit selective for the processing of hand-arm actions conveying vitality forms. It is well known that the main limitation of tractography analysis is the presence of possible high false positive rate. However, the robustness of our results is strongly confirmed by previous tract-tracing studies on macaque [Mesulam and Mufson (1982); Vogt and Pandya (1987)]. These studies showed that in the monkey the sector homolog of the human middle cingulate cortex is tightly connected with the dorso-central part of the insula, known to be involved in modulating hand movements [Jezzini et al. (2012, 2015)]. These data indicate that the cingulo-insular circuit here described appears to be well conserved throughout the evolution of primates.

Although the present study shows that DCI and MCC are both involved in encoding vitality forms perception and expression, results indicate that the MCC is more strongly modulated during the execution of vitality forms compared to DCI, which is instead more significantly modulated during their observation. This finding suggests that these two areas can have a partial different functional role, in line with the existing literature.

Craig (2002) described the whole insula as a sensory “interoceptive cortex” that integrates homeostatic, visceral, nociceptive, and somatosensory inputs, through which a representa-

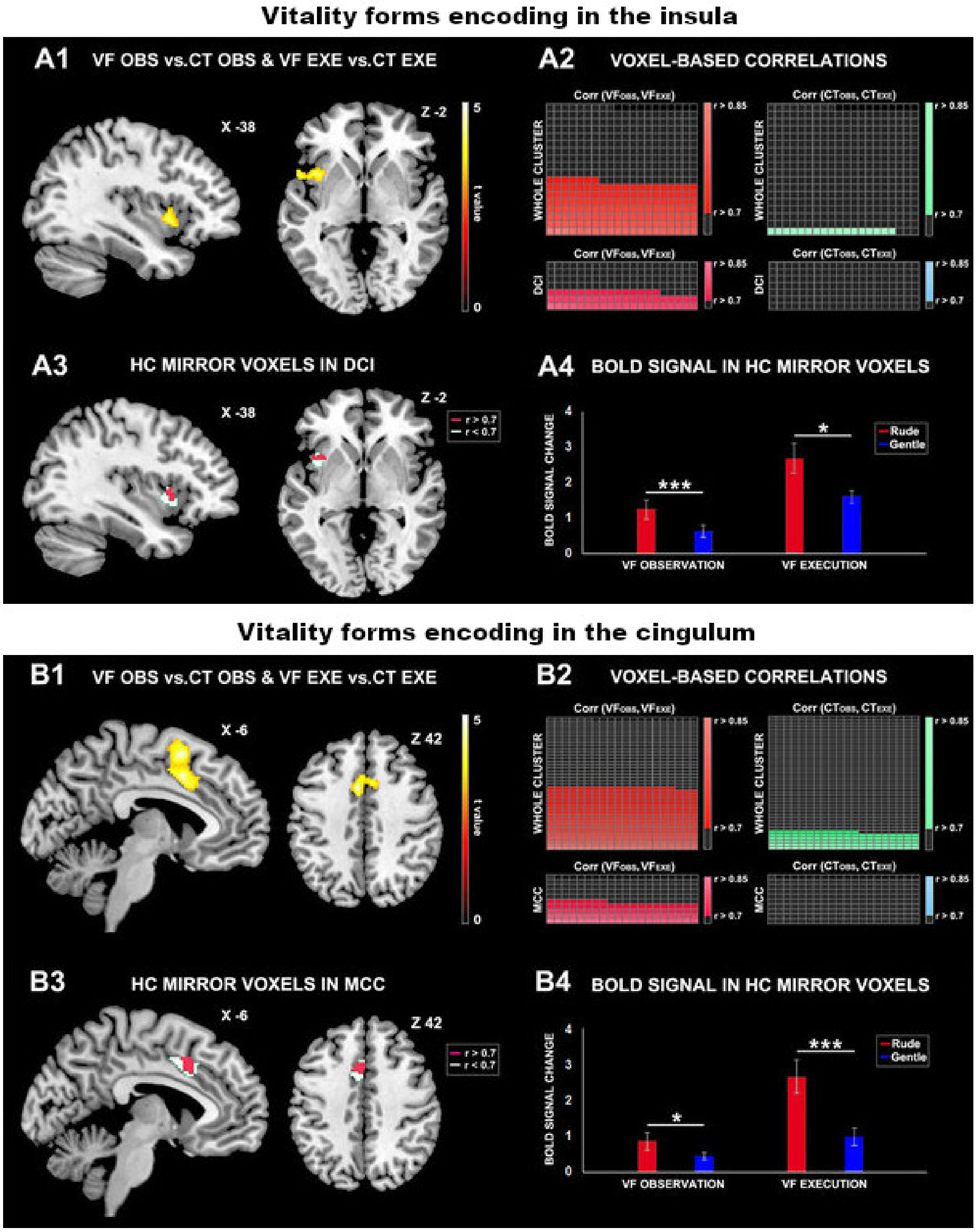


Figure 3.5 Activation of the left insula (A1) and cingulum (B1) resulting from the conjunction analysis of VF OBS vs. CT OBS and VF EXE vs. CT EXE contrast, rendered on a standard MNI brain template (PFWE < 0.05 at cluster level). Maps of voxels showing a high correlated BOLD activity ($r > 0.7$) during the observation and execution of actions with vitality forms (hot colour) or control actions (cold colour) located in the insula (whole cluster, orange colour, A2 Top) and cingulum (whole cluster, orange colour, B2 Top). Voxels located in the DCI (A3) and MCC (B3) showing a high correlated BOLD activity ($r > 0.7$) during VF OBS and VF EXE (HC mirror voxels). Signal changes in the DCI (A4) and MCC (B4) during the processing of gentle and rude vitality forms. Asterisks indicate significant differences at $^{*}p <= 0.05$ and $^{***}p <= 0.001$. Figure adapted from Di Cesare et al. (2021a).

tion of the internal state of the body is generated. Additionally, according to Kurth et al. (2010), the insula can be subdivided in four main sectors: the sensorimotor, socioemotional, olfactory–gustatory, and cognitive ones. The DCI is located in the sensorimotor sector of the insula and is connected with the parieto-frontal circuit for reaching/grasping execution and observation [Di Cesare et al. (2019b)] as well as with temporal areas encoding visual and acoustic biological stimuli [Almashaikhi et al. (2014)]. A study on blindsight patients demonstrated a selective insular activation for conscious perception of fearful bodies, highlighting the role of the insula in the processing of emotional aspects of visual stimuli [Tamietto et al. (2015)].

Concerning MCC, previous studies showed that this cingulate sector is characterized by an evident motor scaffold. Indeed, intracortical stimulation of MCC, carried out on epileptic drug-resistant patients, produced several motor acts, including arm, hand, body, and oral movements. Interestingly, its stimulation produced, before actual movements, the urge to move in relation to external contingencies [Caruana et al. (2018)]. On the basis of these findings and those resulting from the fMRI study presented in this Chapter, it can be hypothesized that DCI plays an essential role in encoding and integrating sensory and interoceptive information for generating the vitality form of the agent during action execution, and for encoding those of the observer during action observation. On the other hand, MCC appears to be more involved in generating vitality forms related to external contingencies, especially in the case of rude vitality forms.

Chapter 4

Two distinct networks for encoding goals and forms of action: an effective connectivity study

Results presented in the following chapter are the outcome of a collaboration with the Wellcome Centre for Human Neuroimaging of the University College of London (UCL), where I spent three months for my PhD visiting period abroad. All these findings have been currently submitted to publication in a high impact journal. In this study, I personally contributed to analyzing data by applying methodology skills acquired abroad, interpreting-discussing results and writing the final paper.

4.1 Introduction

As described in Chapter 2, it is well established that the action goal understanding is mediated by a network formed by a set of parietal and frontal brain areas. This network includes areas of the inferior parietal lobe (IPL) and the premotor cortex (PM). In addition, areas of the posterior superior temporal sulcus (pSTS) become active during hand action observation, but not during hand action execution, suggesting that pSTS is a predominantly perceptual region that recognizes observed actions and sends information to the parieto-frontal network [Rizzolatti et al. (2014)]. It is important to note that this network encodes the action content, but does not appear to mediate the other fundamental aspect of actions: vitality forms. Indeed, as presented in the previous Chapter, the observation and execution of actions endowed with vitality forms activate, in addition to the parieto-frontal network, two main areas, the dorso-central insula (DCI) and the middle cingulate cortex (MCC). While the activation of MCC was unexpected, the activation of DCI is a confirmation of several fMRI studies aiming to assess the neural substrates of vitality forms processing. Moreover, as resumed in the discussion of Chapter 3, numerous studies demonstrated that this sector of the insula is connected to the main nodes of the parieto-frontal circuit. Despite the rich literature concerning structural connectivity of DCI with the parieto-frontal network, there are no data regarding their causal role in terms of effective connectivity. An important issue never investigated concerns the direct causal influences of DCI and the parieto-frontal network during action processing.

To fill this gap, the study described in the present Chapter aims to assess the direction of information flow across four nodes (DCI-pSTS-PM-IPL), establishing their causal role during the processing of hand-arm actions endowed with vitality forms. To this purpose, I carried out Dynamic Causal Modelling (DCM) analysis on data obtained in an fMRI experiment consisting of two main tasks: observation (OBS) and execution (EXE). Particulartly actions (observed/executed) were either endowed with a specific vitality form (gentle or rude) or were performed in a neutral way (constant velocity, control condition) i.e. to minimize, or possibly to eliminate, the presence of vitality forms.

DCM is the predominant framework for the investigation of effective connectivity, i.e. the causal influences among brain regions – or more simply the effect of one region on another. To learn this technique from its theoretical background to its implementation, I spent three months at the Wellcome Centre for Human Neuroimaging of the University College of London (UCL) under the supervision of Professor Peter Zeidman, Chair of the Methods Group and developer of DCM. The present Chapter provides methodology and results of this research project.

4.2 Methods

4.2.1 Participants

Twenty-two healthy right-handed volunteers took part in the experiment(twelve females and ten males, mean age = 22.7, SD = 2.4). All participants had normal or corrected-to-normal visual acuity. None of them reported a history of psychiatric or neurological disorders or current use of any psychoactive medications. All of them gave their written informed consent to be subjected to the experimental procedure, which was approved by the Local

Ethics Committee of Parma (552/2020/SPER/UNIPR) in accordance with the Declaration of Helsinki. Dynamic Causal Modelling analyses was applied to all participants.

4.2.2 Experimental Paradigm

The same experimental paradigm used in the fMRI study presented in Chapter 3 was used: a 2 x 2 factorial design with task (Observation, OBS; Execution, EXE) and conditions (Vitality forms, VF; Control, CT) as factors. Some features were modified to adapt the new fMRI acquisition to the subsequent DCM analysis. Firstly, conditions were presented in mini-blocks of 6 video-clips (instead of 4) each lasting 3 s, for a total 18 s for each mini-block. Secondly, jerky actions observed in the control condition (CT OBS) were replaced with constant velocity actions.

4.2.3 Dynamic Causal Modelling

Theoretical Background

In the DCM framework, the rate of change of neural activity of each brain region at a specific point in time ($\dot{z}$) can be expressed as a function of the experimental inputs (u) and the connectivity between and within brain regions [Zeidman et al. (2019a)]. This is approximated by a neuronal state equation:

$$\dot{z} = (A + \textstyle\sum_{j=1}^{m} u_j B^{(j)})z + Cu.$$

The experimental inputs can enter the model by eliciting direct influences on specific regions at the onset of each stimulus (*driving* inputs) or they can change the strength of coupling among regions by up- or down-regulating specific connections (*modulatory* inputs).

The parameter matrices A, B and C describe three kinds of connectivity which underlie the modelled neural dynamics. Specifically, the A-matrix represents the intrinsic or average (baseline) connectivity within the network of brain regions; the B-matrix represents changes in effective connectivity due to the modulatory inputs; the C-matrix represents the rate of change of the neural response due to the driving inputs. Additionally, each brain region is equipped with an inhibitory self-connection which represents its gain or sensitivity to inputs. From a biological point of view, self-connections control the excitatory-inhibitory balance, mediated by the interaction of pyramidal cells and inhibitory interneurons [Bastos et al. (2012)].

VOI selection and time-series extraction

The first step of DCM analysis entailed the identification of regions of interest and corresponding time-series extraction. In the present study, four nodes revealing a significant response to both the observation and execution of actions conveying vitality forms were selected: posterior superior temporal sulcus (pSTS: -50 -58 6), inferior parietal lobe (IPL: -34 -48 48), premotor cortex (PM: -42 -2 40) and dorso-central insula (DCI: -38 10 -2) (see Figure 4.1). These coordinates found at group level (second level analysis of fMRI) were used to identify corresponding individual activation maxima at the first level (p<0.001 uncorrected). When a participant did not show activation at a specific coordinate, we relaxed the p-value (until p<0.05) to find activated voxels at or close to the expected location. Time-series were extracted by computing the principal eigenvariate within a 5 mm radius sphere centered on each individual maximum.

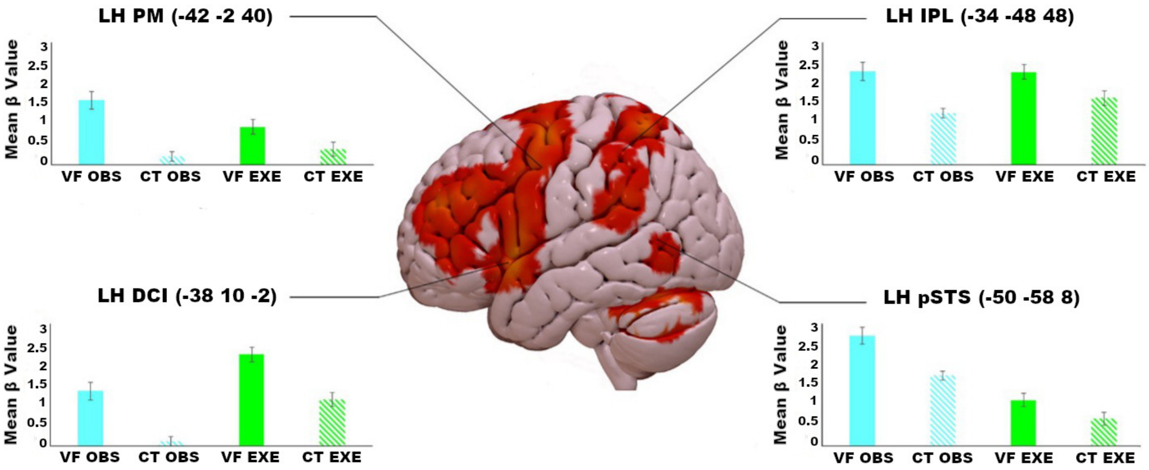


Figure 4.1 Four nodes revealing a significant response (BOLD signal change, mean value) during the observation of vitality forms (VF OBS, light blue bars) vs. control (CT OBS, dashed light blue bars) and the execution of vitality forms (VF EXE, light green bars) vs. control (CT EXE, dashed light green bars): left premotor cortex (LH PM: -42 -2 40), left inferior parietal lobule (LH IPL: -34 -48 46), left dorso-central insula (LH DCI: -38 10 -2) and left posterior temporal sulcus (pSTS: -50 -58 6). These activations are rendered into a standard MNI brain template (pFWE<0.05 at voxel level)

First level DCM analysis

Having identified the regions of interest, the next step is the configuration of the model by specifying which parameters should be switched on and which should be switched off (fixed at their prior expectation of zero). Particularly, a first step consists in the specification of a model in which all plausible parameters are switched on. For this reason, we will refer to this model as a "full" model [Ou et al. (2022); Zeidman et al. (2019a)]. To decide the configuration of the full model, it's important to consider hypotheses and goal of the study. Based on evidence of anatomical connectivity of the four selected regions of interest, for the present study the matrix A involved bilinear connections between all four ROIs (except for the connection between PM and pSTS) and inhibitory self-connections for each region.

To define the C matrix, i.e. driving inputs of the system, we hypothesized that the observation evoked activity in pSTS first and subsequently propagated to the other areas of the circuit; while for the execution we choose PM as most plausible area for the driving input of execution. Thus, from the GLM regressors, OBS (observation) was defined as driving input entering in pSTS and EXE (execution) as driving input entering in PM.

To test the effect of vitality forms on the effective connectivity of the circuit (both during observation and execution), the GLM regressors VF OBS and VF EXE were used as modulatory inputs. Since we did not have hypotheses on the modulation of the system and we did not want to exclude a priori a connection from the modulation of vitality forms (B-matrix), VF OBS and VF EXE were enabled to modulate all connections, including self-inhibitory ones. After computing this full model for each participant, a cell array of all full DCMs (one row per subject) was estimated, yielding the so-called posterior connectivity parameters and their probabilities. In the DCM pipeline, the estimation is the process of finding the parameters (e.g. connection strengths) that offer the best trade-off between explaining the data and minimizing complexity. Model estimation combines the priors with the observed fMRI data to furnish updated posterior beliefs (i.e. after seeing the data) [Zeidman et al. (2019a)].

Second level analysis with PEB

After the estimation of each subject's full DCM, we took the estimated connectivity parameters from each full model to the group level and ran a Parametric Empirical Bayes (PEB) analysis (second level analysis). The PEB analysis captures the commonalities and differences between participants and returns a score (Free energy: F) for the quality of the group-level model [Zeidman et al. (2019b)]. This free energy score quantifies the trade-off between the accuracy and complexity of the model, where more positive values are better. By comparing the free energies resulting from different PEB models, with different set of parameters switched on and off, one can select the model with greatest free energy or model evidence and thereby find the best explanation for the dataset. In the present study, we were interested in identifying the best explanation for the commonalities across our subjects in terms of changes in effective connectivity due to a modulatory effect of vitality forms during action observation and execution. Starting from the full model, a model space containing reduced models with different configurations of the B-matrix was defined. Specifically, in each reduced model, the modulation effect of vitality forms entering in the connection between two regions was switched off. In total, the model space contained 8 candidate models: a full model, 6 reduced models and a “null” model with no modulation to serve as a baseline (Figure 4.2 B). Bayesian model comparison was used to assess the evidence for each model. In addition, in order to summarize parameters across all models and obtain numerical estimates for each of them, a Bayesian model average (BMA) was computed [Hoeting et al. (1999); Penny et al. (2007a)].

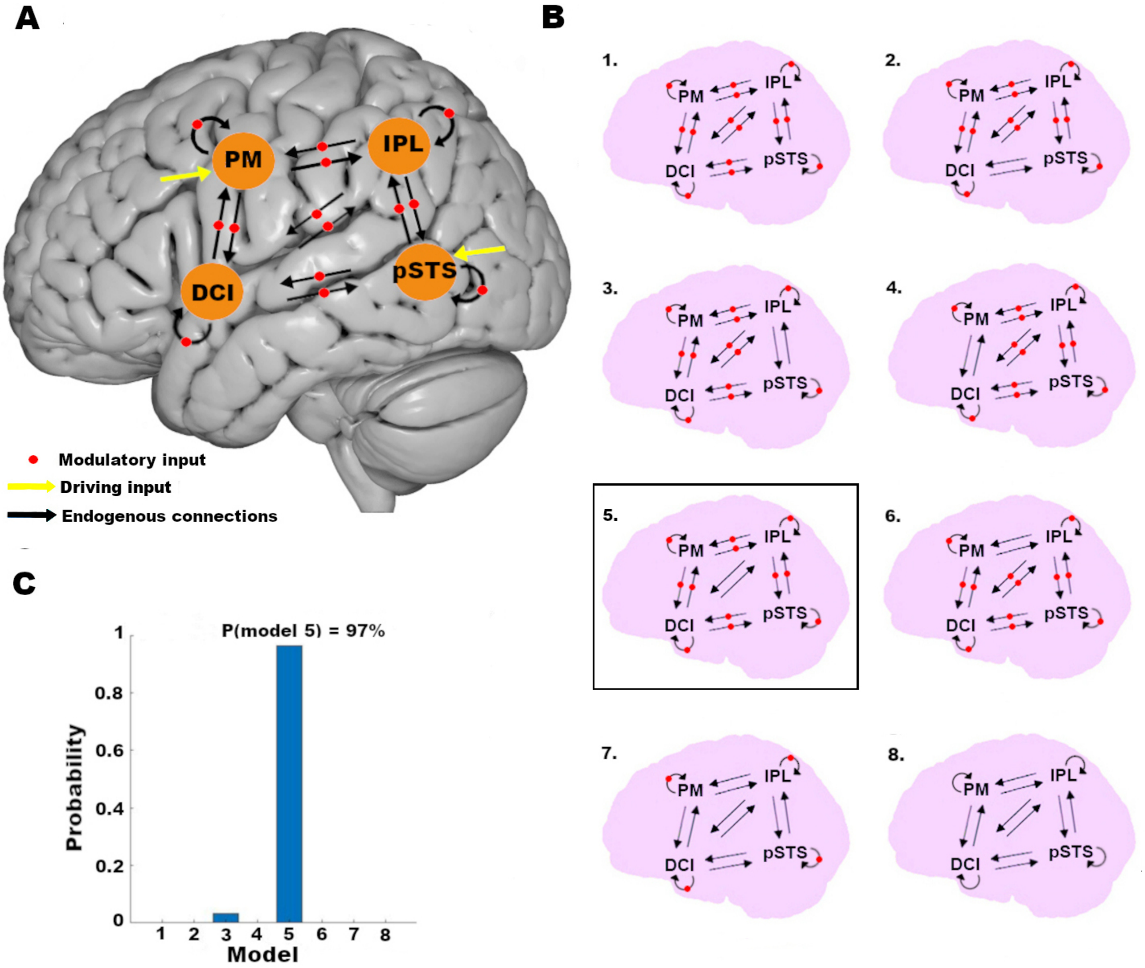


Figure 4.2 The “full” DCM (A). Black arrows represent endogenous connections (A matrix). Yellow arrows represent driving inputs entering in PM (EXE) and pSTS (OBS). Red dots represent modulatory inputs VF EXE and VF OBS (B matrix) which were enabled to modulate all connections. Model space including the “full” model (1), six reduced models (2-7) in which we switched off the modulatory inputs entering in specific connections and a “null” model serving as baseline (8) (B). Model 5, excluding DCI-IPL connectivity, was the winner model with a posterior probability of 97%

.

4.3 Results

Results of the PEB analysis showed that a particular architecture of directed connectivity (model 5) best explained our data, with a posterior probability of 97%, in relation to alternative plausible architectures (Figure 4.2 C). This model precluded connections between DCI and IPL from showing a modulatory effect of vitality forms during both OBS and EXE tasks. Connections (i.e., model parameters) that were modulated by vitality forms were identified by thresholding the Bayesian model average (BMA) at >95% posterior probability. These parameters are shown in Figure 4.3 A1-A2 as connectivity matrices, in which a positive sign (yellow and orange squares) represents a positive modulation while a negative sign (turquoise and blue squares) represents a negative modulation. Self-connections (diagonal values of the matrices) are inhibitory by construction: positive self-connections mean more inhibition and negative self-connections mean disinhibition [Zeidman et al. (2019a,b)].

Considering between-regions connectivity, during the observation of actions conveying vitality forms results revealed a positive modulation of the connections from pSTS to IPL (0.12), from IPL to PM (0.53) and from pSTS to DCI (0.3), from DCI to PM (0.5) (Figure 4.3 A1, 4.4 A1). During the execution of actions conveying vitality forms, results revealed an increased modulation in the directed connections from PM to IPL (0.25) and from PM to DCI (0.41) (Figure 4.3 A2, 4.4 A2). Additionally, results showed a strong "disinhibition effect" via the self-connections of IPL (-1.32) and pSTS (-1.08) during the observation of vitality forms (Figure 4.3 A1) and of PM (-1.14), IPL (-0.57) and DCI (-0.78) during the execution of vitality forms (Figure 4.3 A2).

Finally, to validate its specificity for vitality forms processing during both OBS and EXE tasks, the winning model was tested by verifying how the information flow between its nodes was modulated by the control conditions, comprising the observation (CT OBS) and execution (CT EXE) of neutral actions performed with constant velocity.

Results after BMA thresholding (at >95% posterior probability) revealed that during the observation of control actions there was a positive modulation from pSTS to IPL (0.09) and from IPL to PM (0.22) (Figure 4.3 B1, 4.4 B1). In addition, during the execution of control actions there was a positive modulation from PM to IPL (0.02) and a negative modulation from PM to DCI (-0.13) (Figure 4.3 B2, 4.4 B2). Connectivity parameters involving the DCI did not survive to the BMA thresholding, highlighting the specific involvement of DCI during vitality forms processing only.

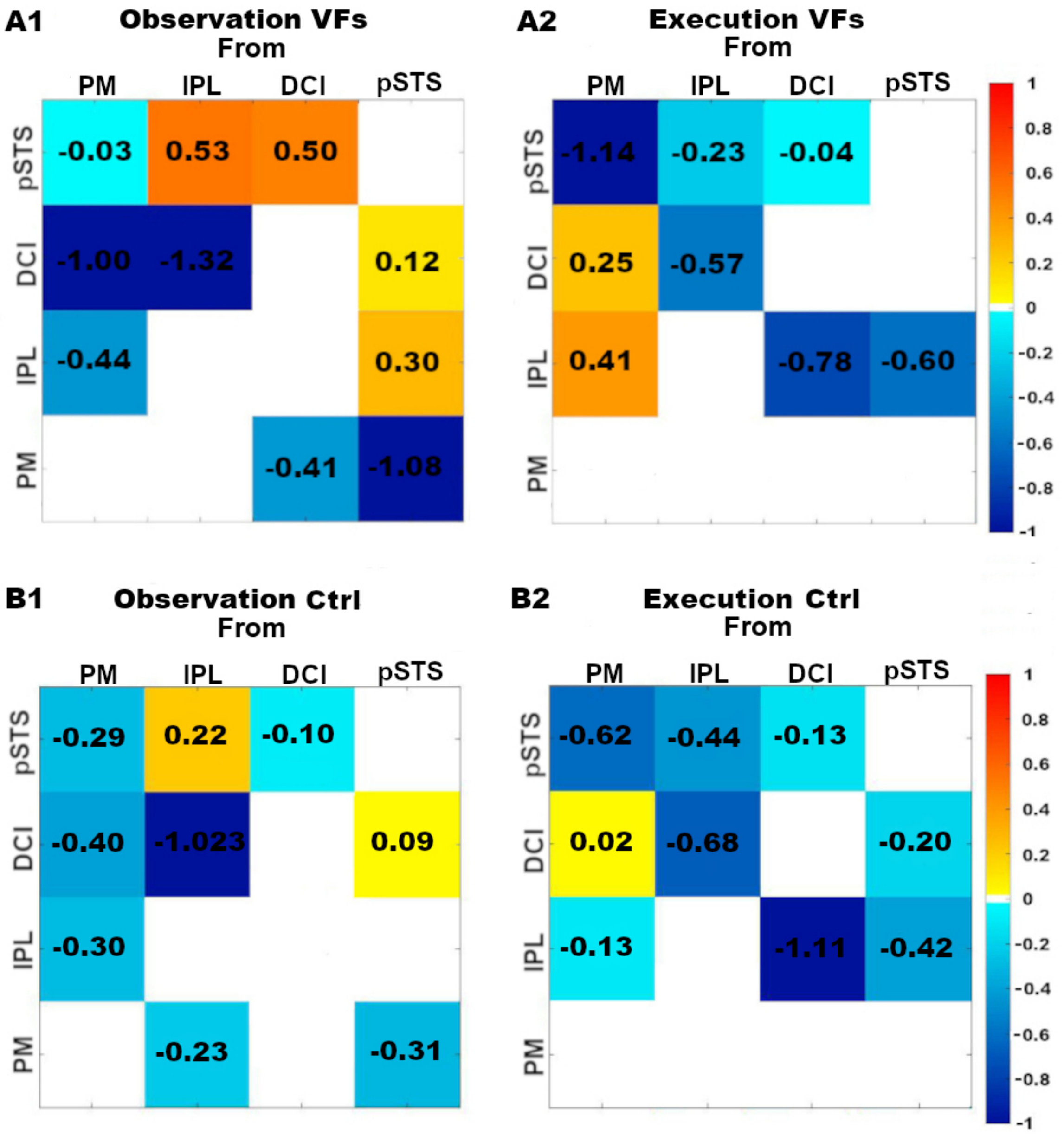


Figure 4.3 Effective connectivity matrices for vitality forms observation (A1), vitality forms execution (A2), control observation (B1) and control execution (B2). For off-diagonal values, connection strengths (Hz) are represented in a scale from yellow to red, if excitatory, and from turquoise to blue, if inhibitory. For leading diagonal values, representative of self-connectivity, which is inhibitory for definition, the colour code is inverted.

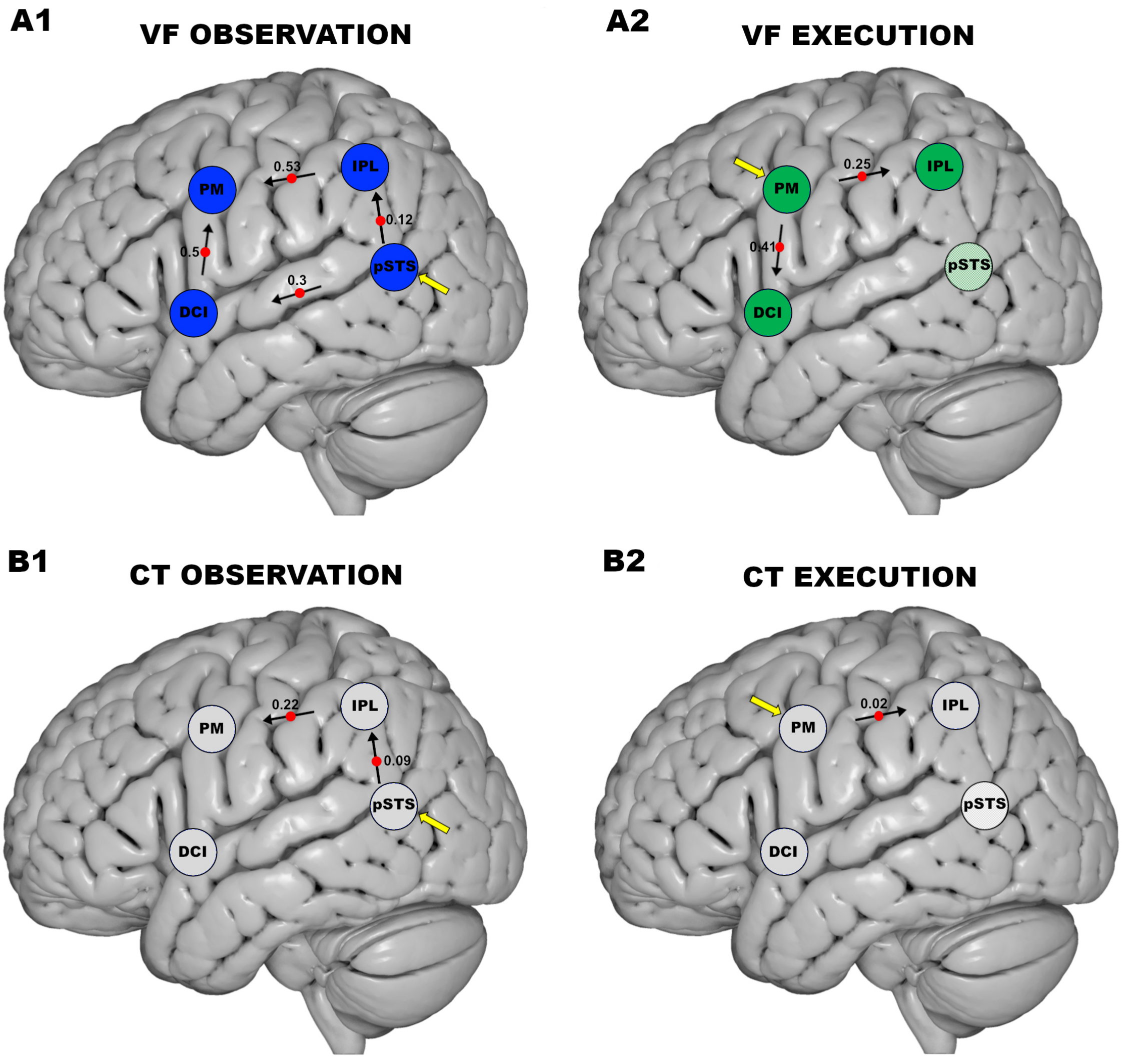


Figure 4.4 Modulatory connection strengths in the winning models surviving at BMA thresholding at 95% posterior probability during vitality forms observation (A1) and execution (A2), control observation (B1) and execution (B2).

4.4 Discussion

In the present study, Dynamic Causal Modelling was used to quantify the directed information flow when processing action goals and vitality forms. For this purpose, an fMRI study employing a classical vitality forms paradigm was carried out. Participants were asked 1) to observe a hand/arm action (observation task); 2) to execute the same action (execution task). Actions were either executed with a specific vitality form (gentle or rude) or in a neutral way i.e. to minimize, or possibly to eliminate, the presence of vitality forms. After conducting standard fMRI analysis, four cortical nodes known to be involved in arm/hand goal directed actions were selected: pSTS, IPL, PM and DCI.

Results show that the observation of actions endowed with vitality forms activates the higher-order visual areas around the posterior part of the superior temporal sulcus (pSTS) from which two streams originate: one reaches IPL and then PM, i.e., the areas forming the classical network involved in the action goal understanding, while the other one, leaving pSTS, reaches DCI and then PM. The presence of this last network supports the view that the stream connecting DCI-PM underlays the action form recognition. It is important to note that, during the observation of actions without vitality forms, a positive modulation of this stream is absent, thus highlighting its specific involvement in vitality forms processing only.

Results relative to the execution of actions endowed with vitality forms show an initial activation of PM, from which two streams depart: one reaches IPL, a region known to be involved, alongside with the premotor motor areas, in the execution of goal-directed actions [Rozzi et al. (2006)], while the other reaches DCI.

The activation of PM during the processing of actions may be related to its encoding of the physical properties of actions [Filimon et al. (2007)]. This is in line with results of Di Dio et al. (2013) demonstrating that, in humans, the premotor and parietal areas are involved in the encoding of reaching movements performed with biological motion. Furthermore, the same authors demonstrated that the observation of different velocities produced an increase of parieto-frontal network activity indicating its specific involvement in the velocity processing. Moreover, Casile et al. (2007) showed that the observation of movements complying with human kinematic laws of motion produced the activation of the dorsal premotor cortex (PMd) extending to the ventral sector (PMv).

Taking together, findings provided in the present chapter clearly indicate that, during the observation and execution of goal directed actions, two streams arise from visual and motor areas: a dorsal stream encoding the action goal and a ventral stream encoding the action vitality form. While the processing of visual information of vitality forms during action observation was in line with the hypothesis already proposed [Di Cesare et al. (2020a)], the modulation effect of PM towards DCI during the execution of vitality forms was unexpected. Indeed, it has been previously suggested that, when an action is executed with a specific vitality form, DCI modulates the parieto-frontal network on the basis of the internal/ affective state of the agent. In this regard a possible limitation should be remarked: during the action execution task, participants had to act in a rude or gentle way after reading an instruction. Thus, it is rather unlikely that, after reading each written instruction (act gently or act rudely), they naturally enter into a specific affective state. Instead, it is more likely that they performed the action with the instructed vitality form using a cognitive command, most probably prefrontal, instead of actually entering into the instructed emotional state. The subsequent modulation of DCI possibly furnishes an affective context for the executed action. These considerations suggest that in PM there is a “repertoire” of kinematically distinct motor acts that may be selected according to the intention on how to behave (e.g., positively or negatively), regardless of the affective state of the agent. This formulation admits a role for the DCI in selecting from the premotor repertoire. This follows because an increase in the effective connectivity from PM to DCI during vitality forms execution means that the DCI responses depend upon the (affective) instructional set, and thereby influence action selection in the PM via reciprocal connections (that do not show a vitality forms modulation). On this view, prefrontal projections may be responsible for instantiating a cognitive set by changing the responses of DCI to PM afferents and, via recurrent connectivity, action selection in the premotor cortex. These findings open a new important perspective regarding the execution of actions conveying vitality forms. It appears that vitality forms could have a “cognitive” or an “affective” origin. The cognitive vitality forms are planned at the onset of an action in the frontal lobe and subsequently activate PM and then DCI. In this way, the motor action acquires an affective component. In contrast, the affective vitality forms originate in DCI and possibly, in some subcortical structures, and modulate the parieto-frontal circuit selecting the appropriate motor act encoded in PM area (gentle action, rude action). As a consequence, a motor act acquires a positive or negative affective aspect. However, future experiments are needed to deeply clarify this point, by generating (affective) vitality forms expression of participants through a mood induction paradigm.

Part II

Behavioural responses

Chapter 5

The influence of vitality forms on action perception and execution

The outcomes presented in the following chapter have been all published in Scientific Reports (Lombardi et al. (2021)). In this study, I personally contributed to performing the experiment, analysing data, interpreting-discussing results and writing the final paper.

5.1 Introduction

Imaging you are interacting with a friend who suddenly asks you to pass an object with a rude tone of voice or by using an aggressive imperative gesture. It is plausible that the negative vitality form used to communicate with you will negatively impact your mood and accordingly modulate your consequent action. On the contrary, if your friend were to ask the same object by using a pleasant and polite tone of voice, accompanied by a delicate gesture, it is likely that the positive vitality form used will positively influence your attitude and subsequent motor response. A first demonstration of this influence comes from the kinematic study carried out by Di Cesare et al. (2017) described in Chapter 2. Results of this experiment showed that participants interacted with the object (giving and taking actions) with a higher velocity and a larger trajectory after a rude request compared to the same interaction after a gentle request. These findings open new issues to be addressed:

Firstly, an insightful point is to investigate whether and how, besides action execution, vitality forms expressed by the agent may also influence the action perception of the receiver. In this regard, previous chapters extensively explained that action perception and execution rely on the same neural correlates and thanks to the mirror mechanism sensory representations of actions performed by others are transformed into one own motor representation. More specifically, results provided in Chapter 3 and Chapter 4 showed the presence of a network selective for vitality forms processing and endowed with mirror properties, being activated during both the observation and the execution of different action forms. Thanks to these findings, it is possible to hypothesize that, in addition to action execution, a request conveying vitality forms may also influence the internal representation of an observed action, modifying the perception of some features such as its estimated duration.

Secondly, Di Cesare et al. (2017) investigated the effect of vitality forms conveyed through visual and vocal stimulations. However, besides actions and speech, social interactions rely on other fundamental communication tools such as touch. Touch is one of the first senses to develop prenatally [Bradley and Mistretta (1975); Marx and Nagy (2015)] and one of the earliest forms of parent-infant communication [Field (2010); Hertenstein (2002)]. In addition to communicating [Hertenstein et al. (2009, 2002)] and eliciting [Suk et al. (2009)] emotions, touch provides an effective means of influencing people's attitudes towards other individuals and their tendency to create bonds and their (pro-)social behaviors [Gallace and Spence (2010)]. Interpersonal touch is used for different purposes. For example, someone can greet us with a delicate or vigorous handshake or ask us something by using a rushed tug or a gentle caress. Interestingly, results of a recent fMRI study showed that a handshake conveying tactile rude/gentle vitality forms produced, relative to a neutral control, a strong activation of the dorso-central insula and the middle cingulate cortex, the two brain areas selectively involved in the processing of actions with vitality forms as described in Chapter 3 [Rizzolatti et al. (2021)]. In this view, it is plausible that vitality forms expressed by an agent through a tactile modality ("physical request") can induce the same influence validated for visual and vocal requests on the motor response of participants.

To tackle these two open issues, the current Chapter describes an experiment aiming to investigate whether and how perceiving different vitality forms can influence cognitive and motor tasks performed by participants. Particularly, participants were stimulated with gentle and rude requests conveyed through two different modalities:

1) physical contact: a robotic manipulandum reproduced a physical request by moving the right arm of participants.

2) voice: participants listened to a vocal request of a human agent.

After these requests, they were asked to perform two tasks:

1) to estimate the end of a passing action observed in a monitor (*action estimation task*).

2) to perform a passing action with the intention to give an object to another person (*action execution task*).

After an extensive description of neural correlates and brain mechanisms underlying vitality forms, the present chapter represents an attempt to investigate their behavioural responses in humans by quantifying their effect in social interactions.

5.2 Methods

5.2.1 Participants

Eighteen healthy right-handed volunteers participated in the present study (12 females and 6 males, mean age = 24.1, SD = 2.7). All participants were native Italian speakers and had normal or correct to normal vision. Nobody reported neurological disorders, physical impairments or current use of psychoactive drugs. All of them gave their written informed consent before participating. This study was approved by the ethical committee of Liguria Region (n.222REG2015) and was carried out according to the principles expressed in the Declaration of Helsinki.

5.2.2 Experimental setup and paradigm

Participants sat in a comfortable chair in front of a monitor, set at a spatial resolution of 1920x1080 pixels, wearing a pair of headphones and holding the handle of Braccio di ferro, a robotic planar manipulandum with a large workspace (80 cm x 40 cm ellipse) and 2 degrees of freedom, with their right hand [Casadio et al. (2006)]. The experiment consisted of two main tasks: action estimation task and action execution task (see Figure 5.1).

In the *action estimation task*, participants were presented with video clips showing the right hand of an actor passing an object (a ball, a bottle, a packet of crackers or a cup) towards another person sitting at the opposite side of a table. Half of the passing actions were performed towards a male actor and the other half towards a female actress. This egocentric perspective was employed to facilitate the motor representation of the observed action, enabling participants to get involved in the action as if the right hand observed belonged to them. By taking inspiration from Aglioti et al. (2008), participants were not presented with the total action duration but only with an initial part (35 % of the entire duration), corresponding to 250 ms for rude actions (total duration 700 ms) and 420 ms for gentle actions (total duration 1200 ms) (see Figure 5.2). The choice of these stimuli comes from results of a preliminary behavioural study in which the same temporal occlusion paradigm was used to validate how participants performed the action estimation task after observing different percentages of duration of the same action (for details see Appendix A

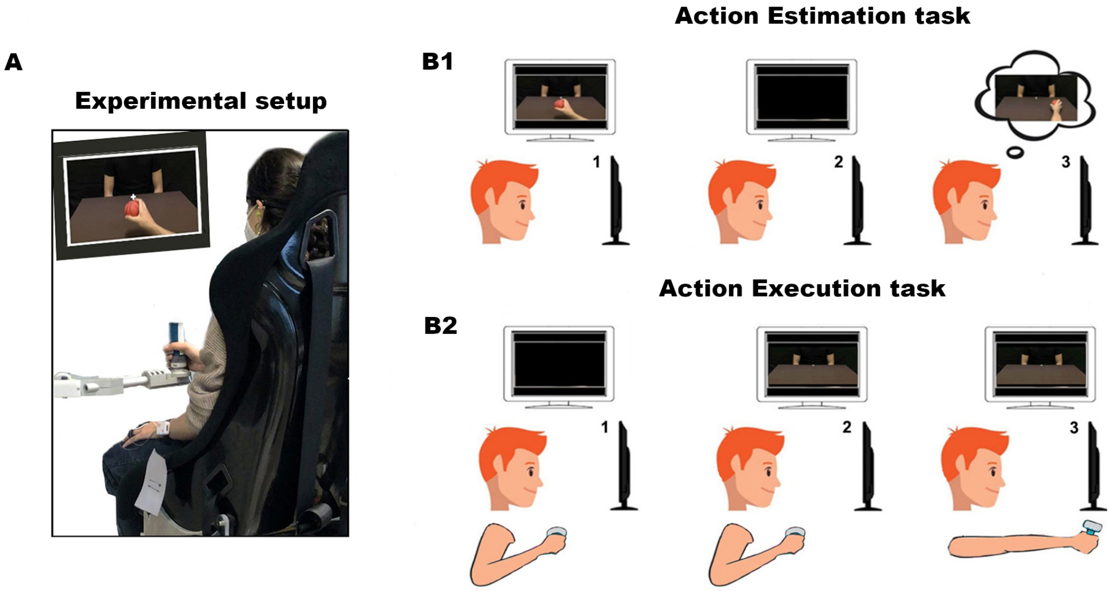


Figure 5.1 Example of a participant during the experiment (A). Action Estimation Task (B1): (1) participants observed the initial part of a passing action, (2) the action was obscured, (3) they continued the action mentally estimating the time of its conclusion. Action Execution Task (B2): (1) starting position, (2) a static image of an actor appeared, (3) participants performed the passage by moving the handle towards the monitor. Figure adapted from Lombardi et al. (2021).

and Di Cesare et al. (2021b)). After having observed the first part of the action, the monitor was obscured with a black screen and participants were required to continue the action mentally and indicate the time of its completion (estimated end) by pressing a little button positioned on the handle of the manipulandum (Figure 5.1 B1).

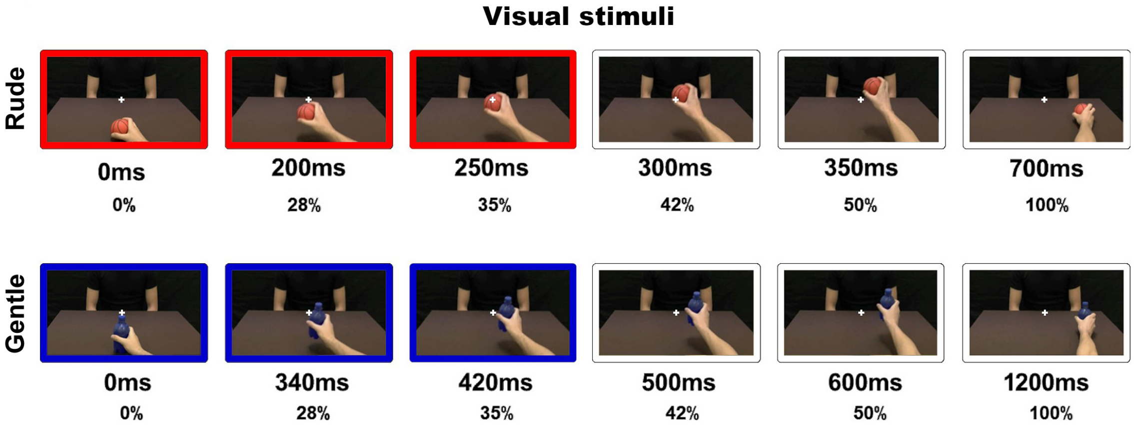


Figure 5.2 During the action estimation task, participants observed only the initial part of the passing actions: 250 ms for rude actions (red) and 420 ms for gentle actions (blue). Figure adapted from Lombardi et al. (2021).

In the *action execution task*, participants were presented with the static image of the same actor or actress and were required to move actively the handle of the robotic manipulandum (motor task) with the intention to pass an object towards the other person (Figure 5.1 B2).
To test the impact of vitality forms on action perception and execution, participants could receive either a physical or vocal stimulation conveying a gentle or rude vitality form before performing a task. We will refer to these stimulations as requests: physical request and vocal request. During the physical request, the robotic manipulandum moved the right arm of participants gently or rudely. During the vocal request, participants listened to the Italian verb “dammi” (English verb “give me”) pronounced gently or rudely by a male or female voice.
In total, the experiment was composed of six runs: two Baseline runs, in which participants simply performed the two tasks without receiving a stimulation before (see Figure 5.3 A1-A2); two Physical Request runs in which participants received the physical request before performing the tasks (see Figure 5.3 B1-B2); two Vocal Request runs in which participant listened to the vocal request before performing the tasks (see Figure 5.3 C1-C2).
The presentation order of the experimental runs was balanced across participants: half participants started with the runs presenting a vocal request followed by the runs presenting a physical request, while the other half started with the runs presenting a physical request

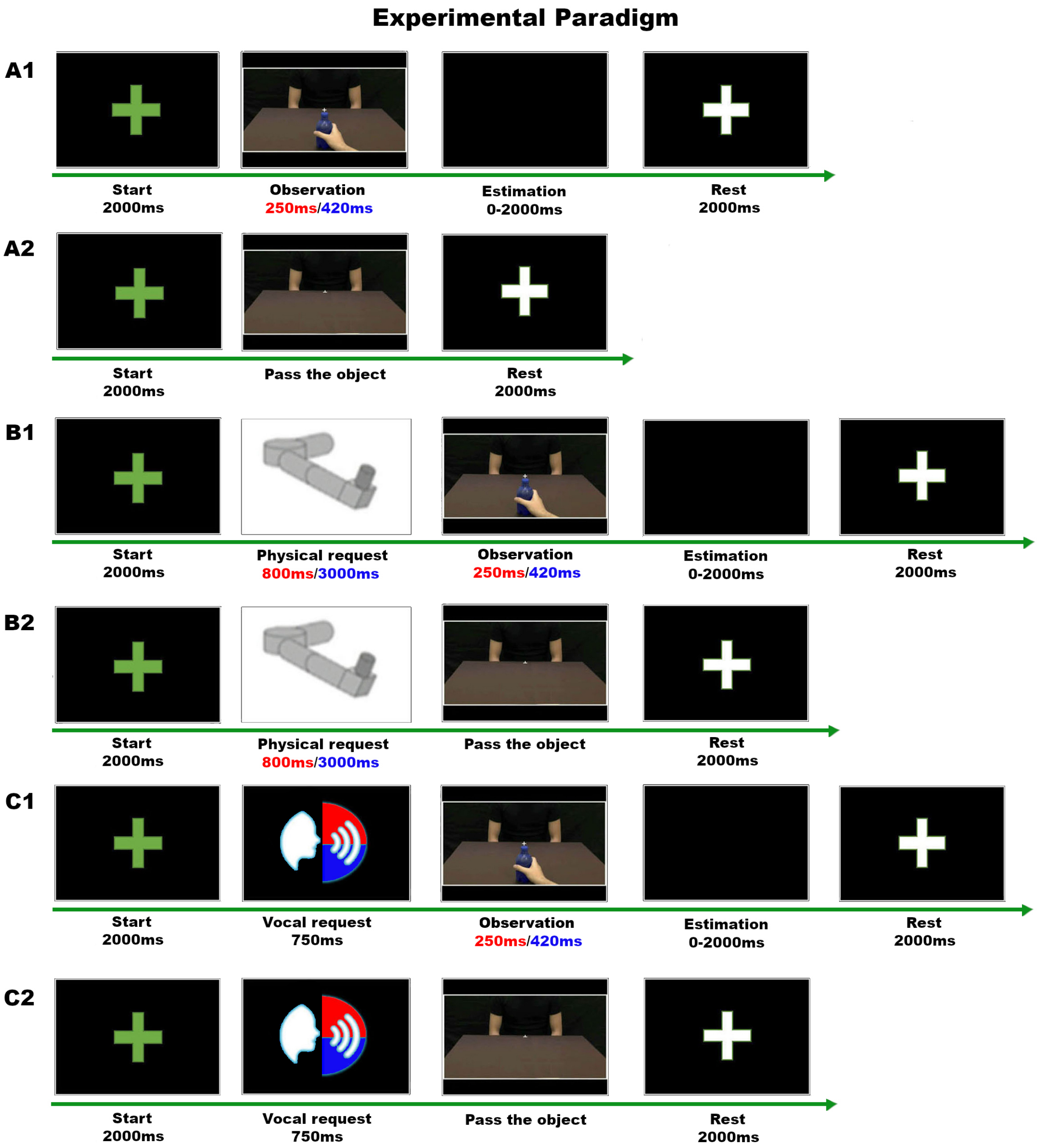


Figure 5.3 Experimental paradigm composed of 6 runs: Baseline runs (A1-A2), Physical Request runs (B1-B2), Vocal Request runs (C1-C2). The panel with the manipulandum icon indicated that participants were stimulated with a physical request before the subsequent task. The panel with the audio icon indicated that participants were stimulated with a vocal request before the subsequent task. Red colour corresponds to rude vitality forms while blue colour corresponds to gentle vitality forms. For each run, the green fixation cross indicated the beginning of a new trial (Start, 2000 ms) and the white fixation cross indicated the rest period before the subsequent trial (Rest, 2000 ms). Figure adapted from Lombardi et al. (2021).

followed by the runs presenting a vocal request. During the Estimation Baseline, 24 stimuli were presented in a random order: 12 rude actions and 12 gentle actions. During the Execution Baseline, participants were required to execute 12 actions by moving the handle towards the monitor. In each of the two Estimation runs, the requests (physical or vocal) and the action subsequently observed could convey the same vitality form (congruent conditions: rude request – rude action; gentle request – gentle action) or opposite vitality forms (incongruent conditions: rude request – gentle action; gentle request – rude action). In particular, in each of these runs, 48 stimuli were randomly presented: 12 rude requests followed by rude actions (congruent conditions), 12 rude requests followed by gentle actions (incongruent conditions), 12 gentle requests followed by gentle actions (congruent conditions) and 12 gentle requests followed by rude actions (incongruent conditions). Finally, in each of the two Execution runs, 12 rude and 12 gentle requests were randomly presented. For each run, a rest period of two seconds was inserted between consecutive trials. It was marked by a white fixation cross on a black screen to keep the attention of participants on the monitor. The white cross turned green before the beginning of a new trial. Before the experiment, participants performed a brief training session. In this phase, they were presented few times with the complete passing actions (rude action: 700 ms, gentle action: 1200 ms) enabling them to familiarize with the original visual stimuli. PsychoPy v3.0 software was used to present video stimuli and to record participants' answers during the action estimation task. Physical requests were instead implemented and controlled through the software environment RT-Lab, integrated with MATLAB/Simulink. RT-Lab included a 100 Hz loop for data storage, which enabled to collect kinematic data during the action execution task.

5.2.3 Physical and vocal requests

As mentioned in the previous section, participants could receive either a physical or vocal request before performing a task. During the physical request, the right arm of participants was moved gently or rudely by the robotic manipulandum. The velocity and trajectory used to implement the physical request derived from previous kinematic recordings in which a human volunteer was asked to move the handle in a gentle or rude way. This procedure allowed to generate a robotic movement which faithfully reproduced a human vitality form (gentle and rude). The final movement consisted in displacements on the horizontal plane starting from the coordinate (0 m, -0.1 m) of the workspace and returning in this starting position at the end of each trial. The gentle request lasted 3000 ms with a maximum displacement of 10 cm in the x-direction, while the rude request lasted 800 ms with a maximum displacement of 22 cm in the x-direction. Additionally, to exclude a phenomenon of adaptation, rude and gentle physical requests were presented in a randomized order and for each vitality form the manipulandum performed three movements with the same velocity but with a small angular shift among them (−10°, 0°, +10°) (see Figure 5.4 A1-A2).

During the vocal request, participants listened to the Italian verb “dammi” (English verb “give me”) pronounced gently or rudely by a male or female voice. Each vocal request was recorded by using a condenser microphone (RODE NT1) placed 30 cm in front of the actors and then digitized with a phantom powered A/D converter module (M-AUDIO M-TRACK). Finally, the audio files were processed with the software COOL EDIT PRO. Both gentle and rude vocal requests lasted 750 ms but differed for parameters such as the pitch and the wave amplitude (see Figure 5.4 B1-B2).

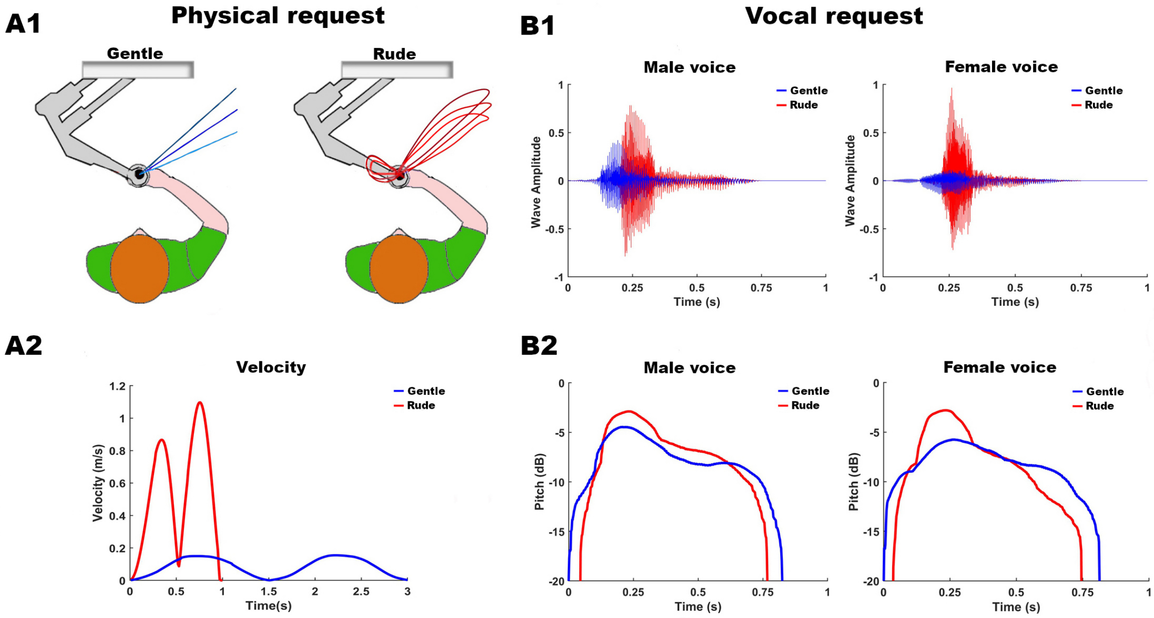


Figure 5.4 Physical request: spatial trajectories (A1) and velocity module (A2) of the movements performed by the manipulandum to provide rude (red lines) or gentle (blue lines) physical requests. Vocal request: Wave amplitude (B1) and pitch (B2) of rude (red) and gentle (blue) vocal requests expressed by the human actor (male voice) and actress (female voice). Figure adapted from Lombardi et al. (2021).

5.2.4 Data analysis

For the action estimation task, participants' responses (estimated action durations) obtained after physical and vocal requests were normalized to the baseline condition:

$$\text{estimated duration after request } (\%) = \frac{\text{estimated duration after request (ms)} * 100}{\text{estimated duration during baseline condition (ms)}}$$

This allowed to obtain percentage values as shown in Figure 5.5 A1-A2. Then, four paired sample t-tests (two for gentle action estimation and two for rude action estimation) were carried out to assess possible differences between congruent and incongruent conditions, after physical (PHY) or vocal (VOC) requests. The significance level was fixed at $p = 0.05$. Before performing statistical analysis, the sphericity of data was verified (Mauchly's test, $p > 0.05$). All variables were normally distributed (Kolmogorov–Smirnov Test, $p > 0.05$).

For the action execution task, action parameters characterizing the motor response of participants after physical and vocal requests were normalized to the baseline condition with the same procedure described above, obtaining percentage values as shown in Figure 5.5 B1-B2:

$$\text{velocity peak after request } (\%) = \frac{\text{velocity peak after request (ms)} * 100}{\text{velocity peak during baseline condition (ms)}}$$

$$\text{distance covered after request } (\%) = \frac{\text{distance covered after request (ms)} * 100}{\text{distance covered during baseline condition (ms)}}$$

Then, four paired sample t-tests (two for action velocity peak and two for distance covered after requests) were carried out to assess possible differences between actions performed after physical (PHY) or vocal (VOC) requests.

5.3 Results

Results of the action estimation task showed a significant difference between congruent and incongruent conditions, for both gentle and rude vitality forms ($p < 0.05$): independently on its modality (physical or vocal), a gentle request increased the duration of the action subsequently observed by participants, while a rude request decreased the estimated duration. More specifically, when participants received a rude request and then observed the initial part of a gentle action (incongruent condition), they anticipated its end compared to the same action presented after a gentle request (congruent condition). In contrast, when they received a gentle request and then observed the initial part of a rude action (incongruent condition), they estimated the action as lasting longer compared to the same action presented after a rude request (congruent condition).

Results of the action execution task showed a significant difference between actions performed after rude and gentle requests, regardless the type of request ($p < 0.001$): independently on its modality (physical or vocal), vitality forms modulated the velocity and the trajectory characterizing the motor response of participants. Specifically, after rude requests, participants executed an action with higher velocity peak and wider trajectory compared to the same action executed after gentle requests. This difference is also highlighted in Figure 5.5 C1-C2 which shows the mean action velocity curves of participants' actions in response to gentle and rude requests (physical and vocal).

5.4 Discussion

The study presented in this Chapter represents the first attempt to quantify the influence of vitality forms on both action perception and execution of individuals during social interactions. Specifically, it aimed to: 1) investigate whether and how vitality forms expressed physically or vocally with gentle and rude vitality forms by an agent may affect participants' responses during a cognitive task (action estimation task); 2) assess how the same requests may modulate the kinematic features characterizing a subsequent action performed by participants (action execution task).

Results of the action estimation task indicated that a gentle request, independently on its modality (through physical contact or vocally), increased the duration of the subsequent action observed by participants. In contrast, a rude request affected the perception of the action subsequently presented, decreasing its perceived duration.

Results of the action execution task indicated that, for both physical and vocal requests, the perception of vitality forms modulated the kinematic parameters of the subsequent action performed by participants. In particular, after a rude request, their action (passing the object) had a higher velocity peak and a wider trajectory. Conversely, after a gentle request, the same action was performed with a lower velocity peak and covered a smaller distance.

Considering the action execution task, it can be argued that this effect can be ascribed to a motor imitation of the vitality form characterizing the request. However, the significant modulation effect of kinematic parameters occurring when participants listened to the vocal request suggests that this effect cannot be solely attributed to an imitation mechanism.

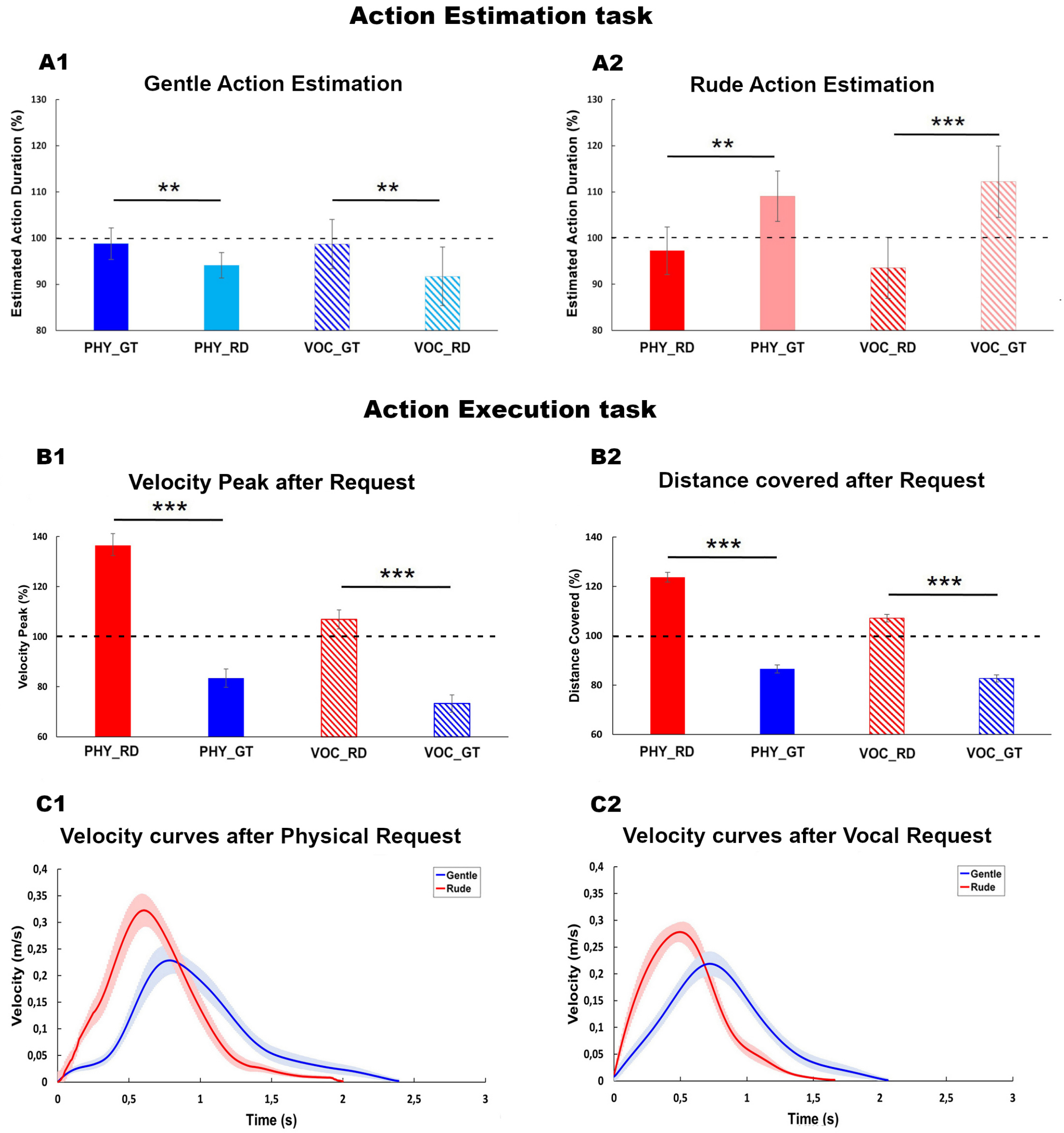


Figure 5.5 Action estimation task: results of gentle (A1) and rude (A2) action estimation. Estimated action durations are normalized to the baseline condition (dotted line in correspondence of 100%). Action execution task: velocity peak (B1) and distance (B2) characterizing the motor response of participants after the request, normalized to the baseline condition (dotted line in correspondence of 100%). Action velocity curves characterizing the passing action performed by participants after a physical (C1) or a vocal (C2) request. Error shadings indicate standard error of the mean. PHY: physical request, VOC: vocal request, GT: gentle vitality form, RD: rude vitality form. Vertical bars represent the standard errors (SE). Horizontal bars indicate statistical significance (**$p < 0.01$, ***$p < 0.001$). Figure adapted from Lombardi et al. (2021).

Moreover, the difference in terms of direction between the physical request (rightward) and the actions observed/executed by participants (frontal direction) contributed to avoid a possible motor imitation.

Another controversial issue one can hypothesize is that the influence of the rude request may be ascribed to a potential arousal effect. Particularly, it is plausible to assert that listening to a rude request may induce the receiver to assume an alert state, making the responses faster. However, this hypothesis is not supported by data obtained during the estimation of actions after gentle requests. Indeed, although participants were asked to estimate the duration of rude actions, if they previously received a gentle request, both physical or vocal, they were induced to estimate the actions as lasting longer. This point is also corroborated by recent findings showing that the effect of a gentle vocal request on rude action estimation was greater than the effect of a rude vocal request on gentle action estimation (see Appendix A at the end of the Chapter).

Taken together, these findings highlight the impact of vitality forms during social interactions, showing how a simple request expressed gently or rudely significantly influences the action perception and execution of the receiver. This effect may be associated to a form of priming, a nonconscious process where exposure to a stimulus alters the response of another stimulus [Langer and Levy-Tzedek (2020); Madhavan and Stoykov (2017)]. For example, in line with the present results, Eizicovits et al. (2018) demonstrated that the movements of a robotic arm primed the subsequent movements performed by participants: they moved significantly slower when interacting with a slow robotic arm compared to when they interacted with a fast-responding one.

While the most part of previous studies aimed to investigate the action goal (what), the present study represents the first attempt to assess the ability in encoding the action form (how) and to remark how much humans are sensitive to this subtle aspect of social communication. Specifically, during the Estimation Baseline run, participants observed the first part of the action and were able to process both its goal (passing action) and its vitality form (rude / gentle) by correctly estimating its time duration. However, when participants perceived a request conveying a specific vitality form before, they were significantly influenced and their cognitive processing of the same action was modulated by that specific vitality form (rude request: estimation of a shorter duration; gentle request: estimation of a longer duration). The same effect occured during the action execution task, in which the kinematic features of the request seemed to be remapped in the motor response of participants influencing the subsequent action. Results provided in Chapter 3 and Chapter 4 represent a neurophysiological explanation of this effect. Indeed, it is plausible that participants processed the action kinematic information of the request and, thanks to the modulation of the parieto-frontal network on the dorso-central insula, they attributed this motor information to an affective component of the action (rude vs gentle vitality form). This information would be remapped in their motor schema enabling to internally represent the associated vitality form and preparing an adequate motor response. This mechanism would allow individuals to obtain a fluidity of interaction that characterizes our everyday encounters with others.

Findings provided by the present study extend the knowledge on the role of the affective states on the affective contagion of others. Previous studies showed how the perception of facial mimicry [Dimberg (1982); Dimberg and Thunberg (2012); Dimberg et al. (2000); Varcin et al. (2019)], body postures [Schmidt et al. (2011)], and vocalizations [Cappella and Planalp (1981); Fujiwara and Daibo (2016)], produced in the receiver an affective convergence supporting the existence of an automatic mechanism selective for the affective contagion. Results of a recent study by Pinilla et al. (2020) indicated that when participants were induced to a positive mood/affective state, they judged both happy and angry faces closer to a positive affective state. In contrast, when participants were induced to a negative mood/affective state, they judged both happy and angry faces closer to a negative affective state. Some limitations of the current study should be considered:

Firstly, the vocal request consisted in one simple imperative action verb “give me,” pronounced by a male or female voice in a rude or gentle way. This limitation may be overcome in the future by reproducing a more realistic dialogue conveying vitality forms.

Secondly, the experiment was carried out on European individuals. Particularly, the physical request was obtained by retargeting kinematic features characterizing gentle and rude actions performed by Italian actors on the robotic manipulandum. Additionally, the vocal request was recorded by the same Italian actors asking them to pronounce the verb "dammi" (English verb "give me") in a gentle or rude way. It is plausible, therefore, that people from different cultures may express positive and negative attitudes towards others in different ways and may employ distinct tones of voice and gestures, reflecting communication styles and cultural norms [Matsumoto (2006)]. For example, David Efron carried out a study on the gestures of Sicilian and Lithuanian Jewish immigrants in New York City, showing that traditional Jews and Italians exhibited distinct sets of gestures. However, as individuals became more assimilated into the American culture, these traditional gestures gradually disappeared [Efron (1941)]. This work was followed by Friesen et al. (1979) who showed cultural differences in emblematic gestures between Japanese, Americans, and New Guineans.

Thirdly, it is plausible that in a realistic scenario the vitality forms effect may have a stronger impact on the receiver being associated to additional social cues of the agent such as facial expressions conveying positive and negative basic emotions.

The next Chapter will focus on overcoming this final limitation by investigating, thanks to the use of the humanoid iCub robot, whether and how positive and negative facial expressions can modify the perception of action vitality forms modulating the effect on human behaviour described in the present chapter.

Appendix A

In completion of data presented in Chapter 5, in this Appendix I provide further results obtained in a preliminary behavioural study which have been published in a journal article for Frontiers in Human Neuroscience (Di Cesare et al. (2021b)). In this study, I personally contributed to writing the paper.

The behavioural study presented in this Appendix was carried out before the implementation descibed in Chapter 5. Particularly, participants were only ask to perform the action estimation task and only the vocal request was used as modality to convey vitality forms.

Visual stimuli consisted of the same video-clips mentioned in the previous chapter: passing actions performed by actors in a rude or gentle way and the same time occlusion paradigm was used (observe only the initial part of the action and estimate the end). However, in this behavioural study, participants could be presented with 28%, 35%, 42% and 50% of the total action duration. Particularly, stimuli could last 200 ms, 250 ms, 300 ms or 350 ms for rude actions and 340 ms, 420 ms, 500 ms or 600 ms for gentle actions.

As shown in Chapter 5, results revealed that the perception of vitality forms expressed by vocal requests influenced the estimation of action duration: a gentle request increased the estimated duration while a rude request decreased it. The effect of vitality forms for different percentages of action presented validated visual stimuli and enabled to chose only one of them for the experiment presented in Chapter 5.

An additional analysis consisted in the evaluation of the overall effect of vocal requests on the subsequent action estimation task. Values obtained from the comparison between congruent and incongruent conditions for the four durations (rude: 200 ms, 250 ms, 300 ms, 350 ms, see Figure A.1 A; gentle: 340 ms, 420 ms, 500 ms, 600 ms, see Figure A.1 B) were averaged. Then a paired sample t-test was carried out to assess possible differences between the effect of gentle request on rude action estimation (red bar, Figure A.1 C) and the effect of rude request on gentle action estimation (blue bar, Figure A.1 C). Results revealed that the effect of gentle vocal request on rude action estimation was significantly greater than the effect of rude vocal request on gentle action estimation (t(29) = 5.2, p = 0.001). As mentioned in the limitations section of Chapter 5, this result excludes that the influence of vitality forms on action estimation was merely due to an arousal effect, because it was easier to slow down the response of participants than to speed it up.

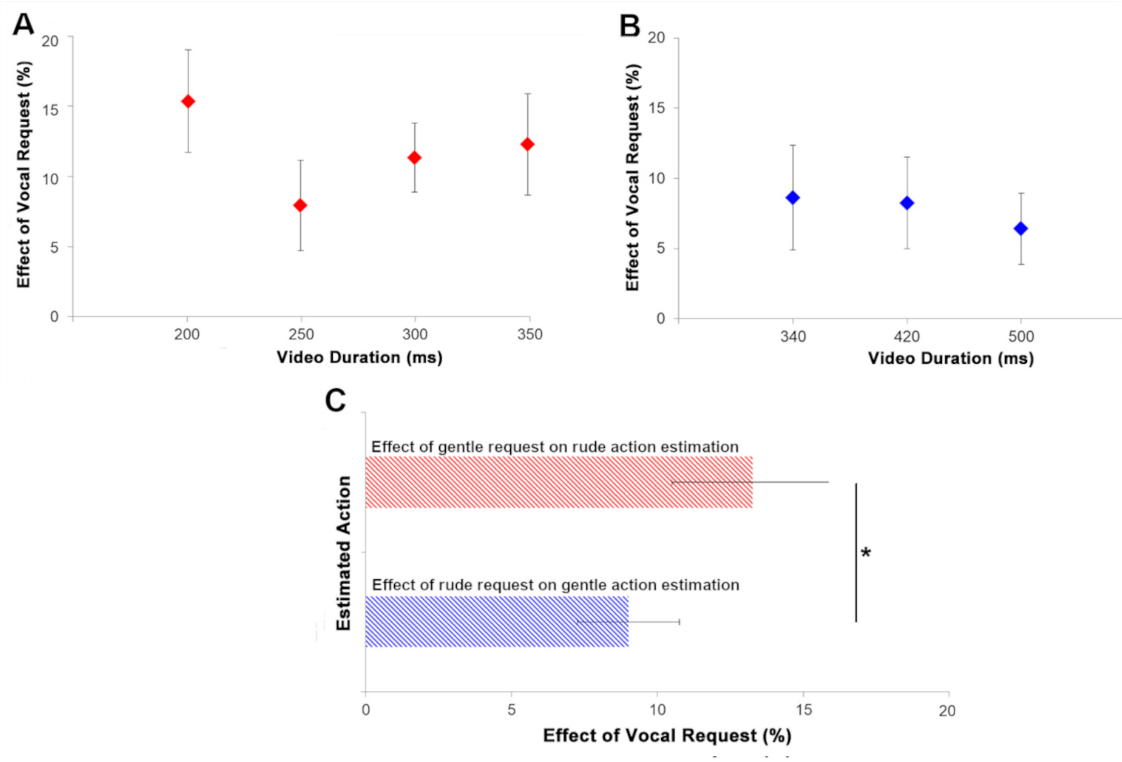


Figure A.1 (A) Effect of gentle request on rude action estimation. (B) Effect of rude request on gentle action estimation. (C) Overall effect showing that A is significantly greater than B. Figure adapted from Di Cesare et al. (2021b).

Part III

From human-human to human-robot interactions

Chapter 6

Humanoid facial expressions as a tool to study human behaviour

The outcomes presented in the following chapter have been all published in Scientific Reports (Lombardi et al. (2024)). In this study, I personally contributed to performing the experiment, analysing data, interpreting-discussing results and writing the final paper.

6.1 Introduction

Results described in the previous Chapter showed that vitality forms expressed by an agent influence the action perception (estimated action duration) and execution (kinematic parameters) of the receiver. It is important to highlight that, in order to focus the attention of participants only on the vitality form conveyed through the physical/vocal requests, additional social cues expressed by the agent were excluded. Also in fMRI studies aiming to study vitality forms conveyed through actions, other emotional components such as the actors' faces were cut from experimental stimuli. However, it is noteworthy that during social interactions, among non-verbal explicit signals, facial expressions represent one of the richest and most powerful resources from which an observer can quickly and easily make inferences about emotional states [Darwin (1872); Ekman (1993); Ekman et al. (1972); Tomkins (1993)], physical health [Jones et al. (2012)], and personality traits [Willis and Todorov (2006)] of the agent.
From a neuroanatomical perspective, it has been recently suggested that the temporal cortex hosts two separate routes for processing static and dynamic faces [Bernstein and Yovel (2015); Furl et al. (2012)]. The ventral stream for faces, including the areas V2-V4 and inferotemporal region, is involved in the recognition of the identity of individuals, while

the dorsal stream for faces, constituted by visual motion area MT, STS, eventually reaching pACC, anterior insula and the amygdala, is considered to be responsible for emotions recognition in healthy subjects.

From a behavioural point of view, emotional contagion theory has been used to explain how a facial expression conveying an emotion affects human behaviour. In particular, people react to facial expressions seen in others tending to mimic them in their own face [Hatfield et al. (1993)]. Additionally, some studies showed that the detection of an emotional expression induces in the observer a corresponding emotional state, thus considering the emotional contagion effect as an initial marker of affective, instead of mere motor-mimetic, reactions [de Gelder et al. (2004); Moody et al. (2007)].

Starting from these findings, an intriguing open issue to address is whether and how different facial expressions of the agent, when associated to gentle and rude actions, can modify the perception of action vitality forms and consequently modulate the response of the receiver. For example, it is plausible that a gentle request, always perceived as communicating a positive attitude of the agent, could have a different influence on participant' responses when associated to a negative facial expression.

To investigate this effect, the study presented in this Chapter proposes the iCub robot as agent of the interaction. Indeed, by using the iCub robot we leveraged on the possibility to manipulate positive and negative facial expressions with action vitality forms, obtaining congruent and incongruent experimental conditions. The same procedure would have been challenging and non-ecological by using human actors, due to the difficulty in naturally dissociating the facial expression from the action (e.g to assume a gentle attitude towards others but performing the associated gentle action with an angry face).

The present experiment was composed of:

1) a preliminary behavioural study, used to validate experimental stimuli and to qualitatively investigate the impact of different facial expressions on the perception of action vitality forms conveyed by the iCub robot

2) a main kinematic experiment, used to quantify the effect of the iCub robot request on participants' motor response

3) a final behavioural study, carried out after the kinematic experiment, in which the same group of participants was required to describe their attitude towards the iCub robot.

6.2 Methods

6.2.1 Participants

Eighteen healthy right-handed volunteers (11 females and 7 males, mean age = 24.28, SD = 2.42) took part in the kinematic experiment. The sample size was defined on the basis of results of an “a priori” power analysis computed with GPower 3.1 [Parameters: effect size f = 0.35; err prob = 0.05; power (1- err prob) = 0.9]. The output of this analysis revealed that a sample size of 16 subjects is sufficient to evidence an interaction effect between the two experimental factors (2 Facial Expressions × 2 Action Vitality Forms). All participants had normal or corrected to normal vision and none of them reported neurological or cognitive disorders. The study received approval by the ethical committee of Liguria Region (n.222REG2015) and was carried out according to the principles expressed in the Declaration of Helsinki. All participants provided written informed consent.

6.2.2 Experimental stimuli

In the present study, the iCub robot represented the agent of the interaction with participants. The iCub platform is a 53 degree-of-freedom humanoid robot of the same size as a three or four years-old child [Metta et al. (2010)], equipped with multiple sensors, including force/torque sensors, encoders in all its joints and eye cameras [Fischer et al. (2018)]. These features allow for an understanding of its body configuration, motor skills and also an ability to show facial expressions, enabling it an ideal platform for human—robot interactions studies.

Experimental stimuli consisted of video-clips showing the iCub robot performing a giving request towards participants: the iCub robot moved its right arm gently or rudely with the palm upward inviting participants to give it a little ball. These actions have been generated by retargeting the kinematic data recorded from a trained actor, with anthropometric measures similar to those of the robot, who performed actions gently or rudely towards an object. This procedure, already validated in other studies, allowed to replicate human actions, enabling the iCub robot to perform the same requests with different vitality forms [Di Cesare et al. (2020b); Vannucci et al. (2018)].

By associating two different action vitality forms (gentle or rude) and two different facial expressions (happy or angry), a total of four conditions were created (Figure 6.1): two congruent conditions in which facial expression and vitality form were characterized by same valence (positive: gentle action and happy facial expression; negative: rude action and angry facial expression) and two incongruent conditions in which facial expression and vitality form were characterized by different valence (rude action and happy facial expression; gentle action and angry facial expression).

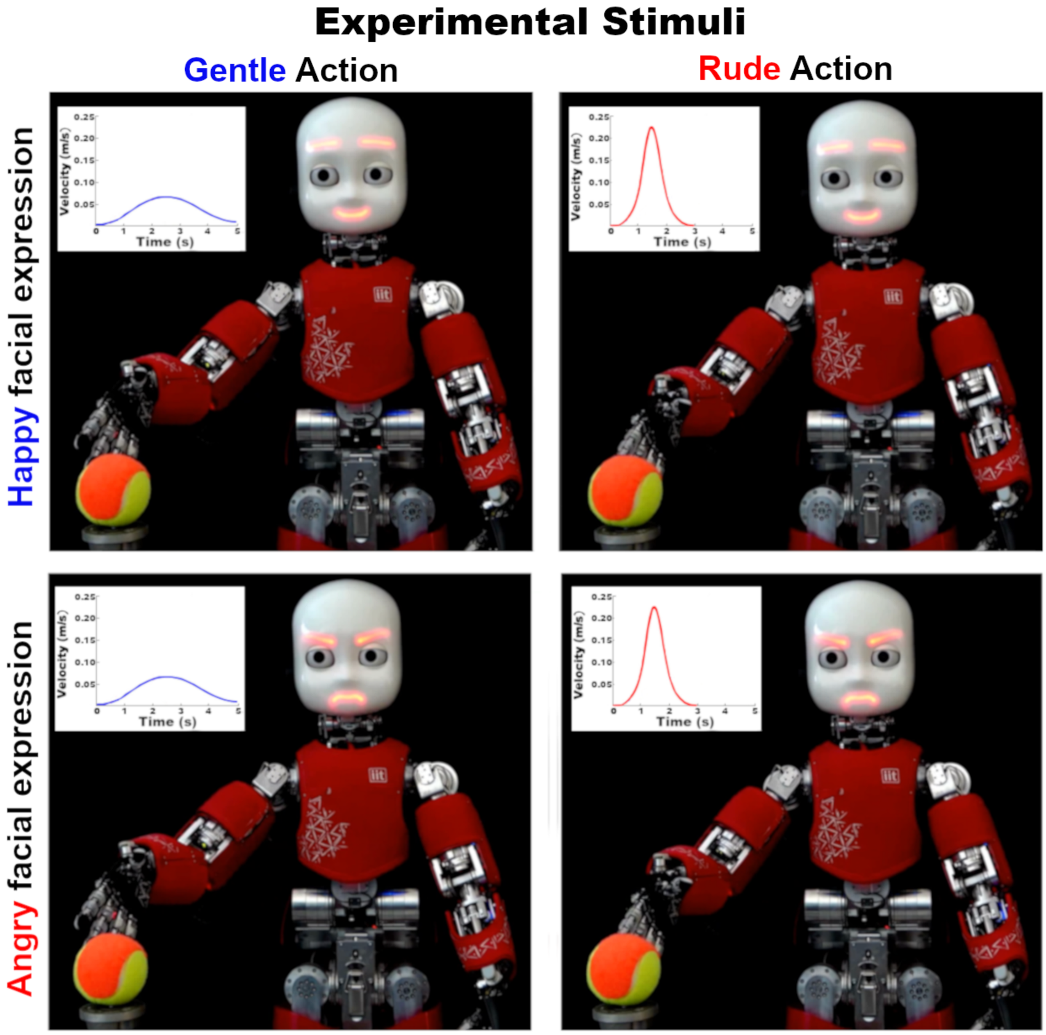


Figure 6.1 Visual stimuli representing the giving request of the iCub robot performed gently (blue curves) or rudely (red curves) with a happy or an angry facial expression. In total 4 conditions were presented: two *congruent* conditions (gentle action-happy facial expression; rude action-angry facial expression) and two *incongruent* conditions (gentle action-angry facial expression; rude action-happy facial expression). Figure adapted from Lombardi et al. (2024).

6.2.3 Experimental setup and paradigm

During the kinematic experiment, participants sat comfortably in front of a table, keeping their right hand with the thumb and index finger set in a pinching position (starting position). This starting position was located 16 cm to the right of the participant's mid-sagittal plane and 25 cm diagonally from the centre of a circled target on which a little ball was placed. Two coloured targets (yellow and orange) were placed 20 cm from the centre of the little ball. The monitor of a PC (25 inch) was positioned closed to the coloured targets in front of participants. Participants were required to observe video-clips representing the iCub robot performing a rude or gentle arm action towards them (giving request) with an happy or angry facial expression. After the request, participants were asked to grasp the little ball and place it on a target with the intention to pass it to the iCub robot presented in the monitor. Before each experimental trial, an instruction concerning the colour (yellow or orange) of the target appeared on the monitor. After their action (reaching and passing), participants returned with the right hand in the starting position. Finally, an instruction on the monitor asked them to replace the little ball on the initial position with their left hand. Between two consecutive trials, participants were simply asked to fix a white cross on a black screen (rest period).
The experiment was composed of two runs, each lasting about 5.30 minutes and consisting of 20 experimental trials presented in a randomized order. In total, 40 stimuli were shown: 20 gentle requests (10 with happy facial expression and 10 with angry facial expression) and 20 rude requests (10 with happy facial expression and 10 with angry facial expression). The experiment was characterized by a 2x2 factorial design, with VITALITY (gentle and rude) and FACIAL EXPRESSION (happy and angry) as factors of interest.

6.2.4 Data recording

The Optitrack system V12O Trio, consisting of a self-contained and factory calibrated tracking bar positioned over the setup, was used to acquire kinematic data. A laptop containing E-prime software was connected to the monitor positioned in front of participants and used for visual stimuli presentation. Kinematic data recording and stimuli presentation were synchronized thanks to an external sync box (Brain Products GmbH) connected between the tracking bar and the laptop. Three passive markers were placed on the right hand of participants. The first marker was positioned on the wrist and was used as reference marker for the extraction of all principal kinematic data. The second and third markers were positioned on the thumb and index fingers nails respectively, used to calculate the maximum hand aperture during the reaching phase and to reconstruct possible gaps in the motion tracking recording

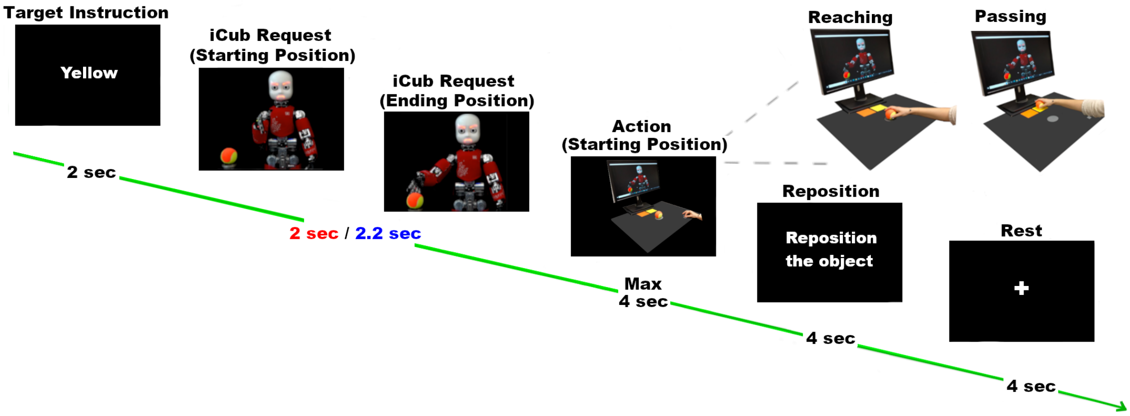


Figure 6.2 After the instrunction of the target colour (Target Instruction, 2 s), participants observed the iCub robot request (gentle: 2.2 s, rude: 2 s) and performed the action (reaching and passing, max 4 s). Then, they were asked to place the ball in the starting position with their left hand (Reposition, 4 s) and to wait the new trial while observing a white fixation cross on a black screen (Rest, 4 s). Figure adapted from Lombardi et al. (2024).

of the first marker. Kinematic data were first pre-processed in the edit layout of Motive 2.3.4 software and subsequently analysed with MATLAB (R2020b).

6.2.5 Data analysis

Each motor response was divided in two phases of interest: the *reaching* phase, during which participants reached the little ball and grasped it, and the *passing* phase, during which participants moved the little ball towards the monitor and positioned it on the requested target. For both phases, specific kinematic parameters were extracted: peak velocity, peak acceleration, z-coordinate trajectory (representing how much participants raised their right hand), action phase duration and time to peak velocity. Additionally, the maximum aperture of the right hand during the reaching phase was calculated.

All kinematic data were normalized to the baseline condition, in which participants performed the task without receiving the iCub request before. In order to investigate whether and how facial expressions and action vitality forms of the iCub robot could influence the motor response of participants, data were organized to carry out a General Linear Model (GLM) for each parameter, with VITALITY (gentle and rude) and FACIAL EXPRESSION (happy and angry) as two factors of interest.

In addition to kinematic features, the reaction time, i.e. the time elapsing between the end of the iCub request and the starting movement of participants, was computed. Also in this case, possible differences of reaction times among conditions were assessed by organizing data in a GLM with VITALITY and FACIAL EXPRESSION as factors.

6.3 Results

Kinematic parameters

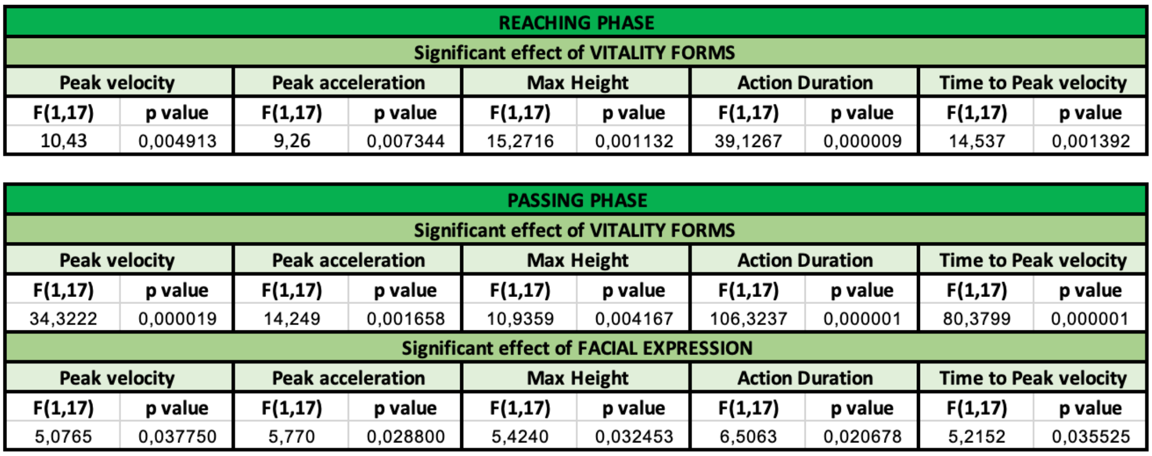
The main effect of VITALITY was significant for the following kinematic parameters for both reaching and passing phases: peak velocity (reaching: F(1,17) = 10.43, $p < 0.01$; passing: F(1,17) = 34.32, $p < 0.001$; Figure 6.4 A1-B1), peak acceleration (reaching: F(1,17) = 9.26, $p < 0.01$; passing: F(1,17) = 14.25, $p = 0.001$; Figure 6.4 A2-B2), max height (reaching: F(1,17) = 15.27, $p = 0.001$; passing: F(1,17) = 10.93, $p < 0.01$; Figure 6.4 A3-B3), action duration (reaching: F(1,17) = 39.127, $p < 0.001$; passing: F(1,17) = 106.32, $p< 0.001$; Figure 6.5 A1-A2), time to peak velocity (reaching: F(1,17) = 14.54, $p < 0.01$; passing: F(1,17) = 80.38, $p < 0.001$; Figure 6.5 B1-B2).

The main effect of FACIAL EXPRESSION was significant for the following kinematic parameters in the passing phase only: peak velocity (F(1,17) = 5.07, $p < 0.05$; Figure 6.4 C1), peak acceleration (F(1,17) = 5.77, $p < 0.05$; Figure 6.4 C2), max height (F(1,17) = 5.42, $p < 0.05$; Figure 6.4 C3), action duration (F(1,17) = 6.50, $p < 0.05$; Figure 6.5 A3), time to peak velocity (F(1,17) = 5.21, $p < 0.05$; Figure 6.5 B3). No interaction effect VITALITY*FACIAL EXPRESSION was found. Analysis of the maximum hand aperture did not reveal any significant main effect of VITALITY and FACIAL EXPRESSION.

The lower part of Figure 6.4 depicts the velocity (Figure 6.4 D1-E1), acceleration (Figure 6.4 D2-E2) and trajectory (z-coordinate; Figure 6.4 D3-E3) curves of participants actions in response to the iCub robot requests performed gently (blue curves) or rudely (red curves) with happy (Figure 6.4 D1-D2-D3) or angry (Figure 6.4 E1-E2-E3) facial expression. For a clearer overview of these results see the table in Figure 6.3.

Reaction time

Results of the reaction time analysis revealed a significant main effect of VITALITY (F(1,17) = 7.60, $p < 0.05$), a significant main effect of FACIAL EXPRESSION (F(1,17) = 12.94, $p < 0.01$) and a significant interaction effect VITALITY*FACIAL EXPRESSION (F(1,17) = 51.04, $p<0.001$, Figure 6.5 C2). Post-hoc analysis (Newman-Keuls correction) revealed a significant difference of reaction times among all conditions except for the comparison of congruent ones (Figure 6.5 C1).



REACHING PHASE									
Significant effect of VITALITY FORMS									
Peak velocity		Peak acceleration		Max Height		Action Duration		Time to Peak velocity	
F(1,17)	p value	F(1,17)	p value	F(1,17)	p value	F(1,17)	p value	F(1,17)	p value
10,43	0,004913	9,26	0,007344	15,2716	0,001132	39,1267	0,000009	14,537	0,001392

PASSING PHASE									
Significant effect of VITALITY FORMS									
Peak velocity		Peak acceleration		Max Height		Action Duration		Time to Peak velocity	
F(1,17)	p value	F(1,17)	p value	F(1,17)	p value	F(1,17)	p value	F(1,17)	p value
34,3222	0,000019	14,249	0,001658	10,9359	0,004167	106,3237	0,000001	80,3799	0,000001
Significant effect of FACIAL EXPRESSION									
Peak velocity		Peak acceleration		Max Height		Action Duration		Time to Peak velocity	
F(1,17)	p value	F(1,17)	p value	F(1,17)	p value	F(1,17)	p value	F(1,17)	p value
5,0765	0,037750	5,770	0,028800	5,4240	0,032453	6,5063	0,020678	5,2152	0,035525

Figure 6.3 Significant effects resulting from statistical analysis of kinematic parameters.

6.4 Discussion

The present study aimed to investigate whether and how positive and negative facial expressions of an agent could modify the perception of their action vitality forms and consequently the motor response of a receiver. To determine this possible effect, congruent and incongruent conditions were tested. Specifically, in the congruent conditions the facial expression and the action vitality form of the agent were characterized by the same valence (positive: gentle action and happy facial expression; negative: rude action and angry facial expression) while in the incongruent conditions the facial expression and the action vitality form of the agent were characterized by opposite valence (rude action and happy facial expression; gentle action and angry facial expression). It is important to remark that the incongruent conditions were made possible thanks to the use of a humanoid such as the iCub robot. In this regard, the use of the iCub robot as interactive and controllable agent to investigate the influence of facial expressions on vitality forms perception during social interactions represents an innovative point of the present study.

Results of the first behavioural study, carried out to validate experimental stimuli, indicated that, independently from the action vitality form observed (gentle or rude), the information coming from the facial expression (happy or angry) guided the choice of participants in the description of the attitude conveyed by the iCub robot (see Appendix B at the end of the Chapter). Particularly, the valence chosen for the description of the iCub robot in incongruent conditions was the one coming from the facial expression. For example, if the iCub robot moved gently but showing an angry facial expression, participants mostly chose the adjective "angry" to describe it. These data suggest that the positive and negative facial expressions

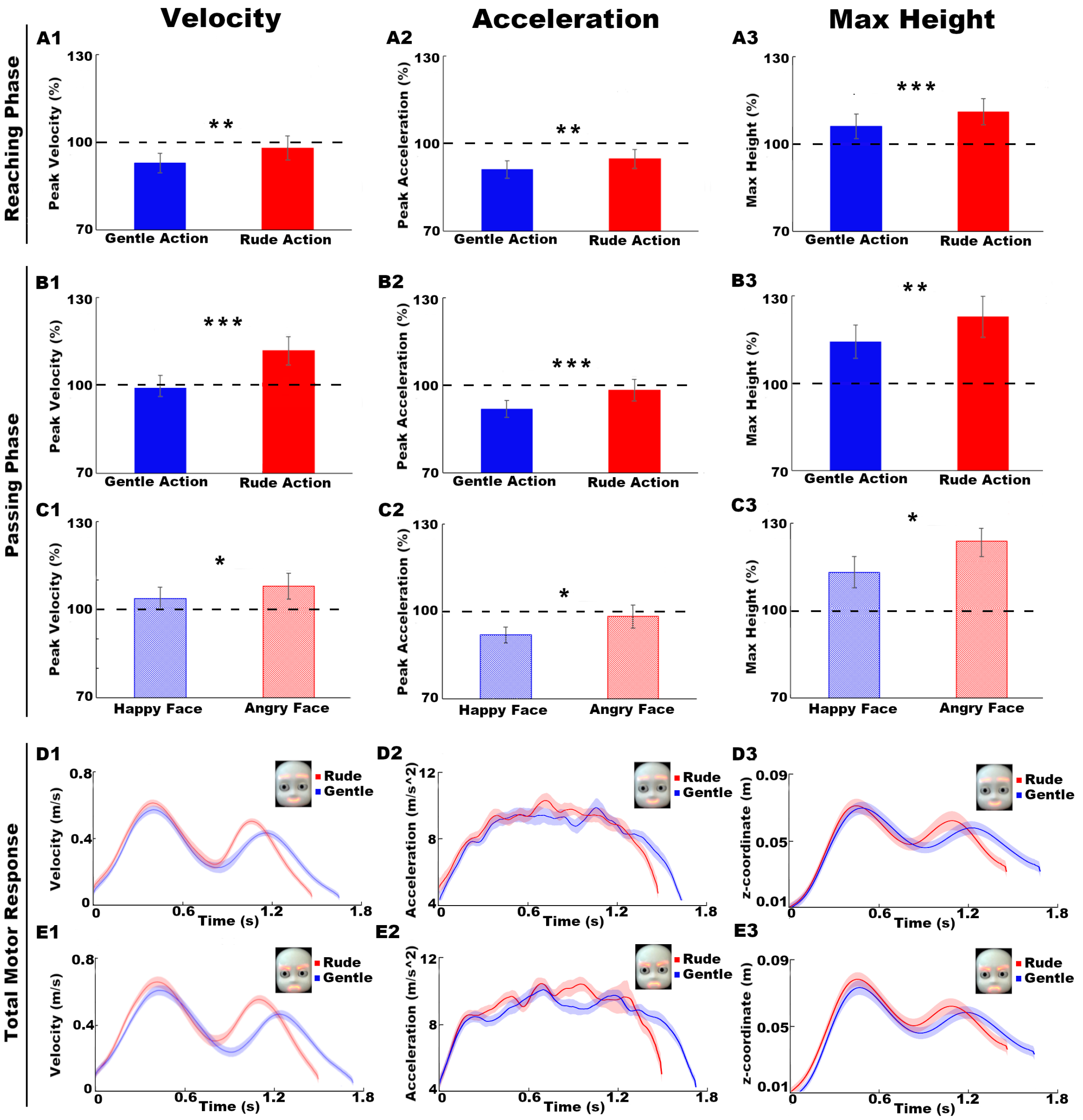


Figure 6.4 Main effect of VITALITY (gentle: blue bars, rude: red bars) on peak velocity, peak acceleration and max height during the reaching (A1-A2-A3) and passing (B1-B2-B3) phases. Main effect of FACIAL EXPRESSION (happy: shaded blue bars, angry: shaded red bars) on peak velocity (C1), peak acceleration (C2) and max height (C3) during the passing phase; The dotted line in correspondence of 100% refers to the baseline condition. Vertical bars represent the standard errors. Horizontal bars indicate statistical significance (*p <= 0.05, **p <= 0.01, ***p <= 0.001). Graphs below show the velocity (left panel), acceleration (middle panel) and z-coordinate (right panel) curves characterizing the total motor response of participants after gentle (blue) and rude (red) requests performed by the iCub robot with a happy (D1-D2-D3) or angry (E1-E2-E3) facial expression. Error shadings indicate standard error of the mean. Figure adapted from Lombardi et al. (2024).

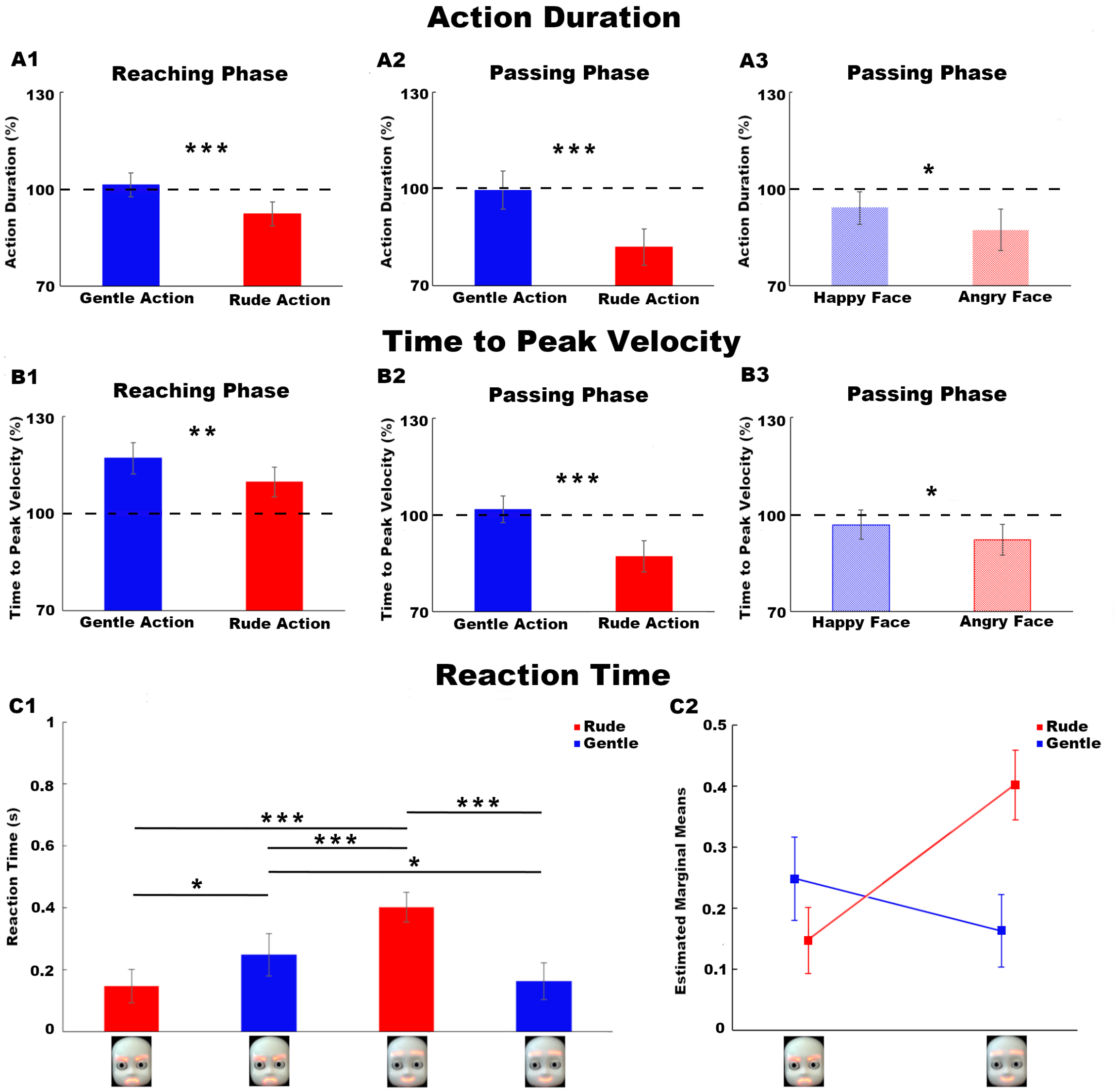


Figure 6.5 Main effect of VITALITY (gentle, blue bars; rude, red bars) on action duration (A1: reaching phase; A2: passing phase) and time to peak velocity (B1: reaching phase; B2: passing phase). Main effect of FACIAL EXPRESSION (happy, shaded blue; angry, shaded red) on action duration (A3) and time to peak velocity (B3) during the passing phase. The dotted line in correspondence of 100% refers to the baseline condition.
Reaction time: Post Hoc analysis (C1) showing significant differences among conditions and significant interaction VITALITY*FACIAL EXPRESSION (C2). Vertical bars represent the standard errors. Horizontal bars indicate statistical significance (*p <= 0.05, **p <= 0.01, ***p <= 0.001). Figure adapted from Lombardi et al. (2024).

modify the perception of action vitality forms.

Results of the kinematic experiment confirmed and deepened this hypothesis, showing that vitality forms expressed by the iCub request significantly influenced the motor response of participants, during both the reaching and passing phases. Most importantly, the facial expression of the iCub robot significantly modulated several kinematic parameters during the passing phase. Specifically, if the iCub robot performed a rude request with a happy facial expression, compared to the same request with an angry facial expression, participants perceived it as communicating a positive attitude. Consequently, they interacted with it by decreasing velocity peak, acceleration peak and maximum height of movement (Figure 6.4) and by increasing action duration and time to peak velocity (Figure 6.5). On the contrary, if the iCub robot performed a gentle request with an angry facial expression, compared to the same request with a happy facial expression, participants perceived it as communicating a negative attitude and the influence on kinematic features reversed. These data suggest that, as hypothesized after the preliminary behavioural study, positive (happy) or negative (angry) facial expressions can modify the perception of the same action vitality form (positive: gentle; negative: rude) and affect the interaction of participants towards the robot.

To further discuss results of the kinematic study, it is important to point out the different features characterizing the reaching and passing phases. On the one hand, the reaching phase consisted of an action directed towards a little ball, thus requiring participants to be mostly focused on the object to correctly grasp it. For this reason, while action vitality forms of the iCub robot modulated kinematic parameters by accordingly increasing or decreasing them, its facial expressions did not produce a significant modulation. On the other hand, the passing phase consisted of an action directed towards the iCub robot, agent of the interaction, thus enabling participants to naturally express their attitude towards it. In this case, besides action vitality forms, also the facial expressions of the iCub robot significantly modulated kinematic parameters. This interpretation finds evidence in several studies aiming to study the effect of different social contexts and intention on arm kinematics. In the late 1980s, Marteniuk et al. (1987) showed that kinematic features of participants actions were modulated when task demands required greater precision, showing that movement production is relatively specific to the constraints of the subsequent task. This effect was also present when participants were required to grasp an object to eat it or move it [Naish et al. (2013)] and grasp an object to lift or insert it into a niche [Ansuini et al. (2006)]. Becchio et al. (2008) demonstrated similar effects for social intentions, i.e. intentions directed towards another person. In this study, participants were required to move an object from a location to another (individual intention condition) or to pass it to a partner (social intention condition). Different kinematics patterns were observed for the individual and social conditions.

Besides motor behaviour, various studies also highlighted the effect of social context on the human cognitive behaviour. For example, it has been shown that the reaction time (i.e the time occurring between a stimulus and the initiation of a motor response) can vary depending on different social conditions [Quesque and Coello (2014); Quesque et al. (2017)]. The present belief is that two important processes occur within this reaction time: the selection of *what* needs to be achieved and *how* [Wong et al. (2015)]. Results of the present kinematic study showed that both action vitality forms and facial expressions of the iCub robot significantly modulated the participants reaction time, i.e. the time occurring between the end of the iCub request and the start of their motor response. The significant interaction between the two factors of interest (VITALITY and FACIAL EXPRESSION) indicates that the reaction time is remarkably modulated by the perception of different action vitality forms and facial expressions. As illustrated in Figure 6.5 C1, participants were more rapid and spontaneous to initiate the motor response in congruent conditions suggesting that, when the facial expression and the action vitality form were characterized by the same valence (positive: gentle action and happy facial expression; negative: rude action and angry facial expression), participants easily processed this information and rapidly associated an attitude to the iCub robot. In contrast, during the incongruent conditions, when the facial expression and the action vitality form were characterized by opposite valence (rude action and happy facial expression; gentle action and angry facial expression), participants spent more time in the cognitive processing of the attitude communicated by the iCub robot, lengthening the time between its request and their motor response.

Results of the final behavioural study, carried out on the same group of participants, showed that, depending on the request, participants differently described their attitude towards the iCub robot. Interestingly, when the iCub robot performed a gentle request with an angry facial expression (incongruent condition), participants stated to be “threatened” by it. This suggests that the negative (angry) facial expression modifies the meaning typically associated to the gentle action, making the iCub robot to be perceived as a commanding agent. Furthermore, when the iCub robot expressed a rude request with a happy facial expression (incongruent condition), participants stated to be “surprised”. In this case, it is plausible that participants felt astonished after viewing a positive (happy) facial expression associated to a rude action, a result which easily explains the longest reaction time observed in the kinematic study during this specific condition.

Some limitations of the current study should be considered:
Firstly, the iCub facial expression did not modulate kinematic parameters during the reaching phase and neither facial expressions neither action vitality forms of the iCub robot influenced the maximum aperture of the grasping hand. The absence of these effects can be due to constraints related to the object used (little ball), which requires a precise grip for a correct movement. Future experiments could overcome this limitation by replacing the little ball with more easily graspable objects, such as a bottle or a cup.
Secondly, the iCub robot is able to convey different basic emotions with its facial expressions (e.g happiness, anger) by simply turning on the led light of eyebrows and mouth. Displaying emotional facial expressions by using predefined target positions of led light does not properly reflect the inherent characteristics in human facial dynamics. Thus, the use of static facial expressions may have reduced the effect on participants' motor response. Emotion recognition ability has been widely studied both by using static stimuli and dynamic ones (neutral faces gradually unfolding into emotional expressions). Since real-life social interactions are characterized by dynamic facial behaviors it had been hypothesized that the motion inherent to facial expressions plays a crucial role in understanding them correctly [Bruce and Young (1976)]. For example, Lazzeri et al. (2018) evaluated whether the expressions performed by a humanoid robot are positively influenced by the dynamic component as it happens in case of human facial expressions by comparing static and dynamic expressions of a human actress and a humanoid robot. Results related to the recognition scores showed that static stimuli were more ambiguous than the dynamic stimuli both for human and robot facial expressions. Interestingly, negative expressions, such as fear and disgust, were more difficult to recognize in comparison with the positive expressions, confirming previous findings stating that positive emotions conveyed through facial expressions may be visually simpler to be recognized than negative ones [Adolphs (2002); Leppanen and Hietanen (2004)]. In this view, studies on implementing dynamic facial expressions of robots are indispensable in enabling more effective (and natural) interactions with humans. Park et al. (2015) proposed a dynamic emotion model as a method of generating more realistic robot facial expressions. Particularly, by using a second-order differential equation to obtain a trajectory between different emotions and simple adjustments of coefficients (such as rise time Tr) rather than using a linear interpolation, the authors built a variety of expression changes (nine dynamics). Moreover, they add facial actions such as sniffling or wailing loudly for sad, laughing aloud for happiness, etc, obtaining several realistic facial expressions in robots in the case of even the same emotion. This is important in relation to results of an fMRI-EMG study carried out during my PhD [Lombardi et al. (2017)] which showed that, besides actions, also

facial expressions can convey basic emotions with different forms. For example, joy can be expressed by a weak smiling face (low intensity) or an overt laughing (high intensity), modulating accordingly the participants' facial reactions (emotional contagion). Taking these findings together, a fascinating perspective would be to endow the iCub robot with the ability to associate different action vitality forms to dynamic facial expressions by creating different level of intensities and led configurations for each basic emotion.

To conclude, the present Chapter showed that, besides action vitality forms, facial expressions represent essential social cues to understand the state of an agent. Particularly, positive and negative facial expressions can modify the perception of gentle and rude actions they are associated with, affecting the motor interaction of individuals towards the agent. Notably, the present study represents the first attempt to use the humanoid iCub robot as novel tool to investigate how positive and negative attitudes conveyed through the association of facial expressions and action vitality forms (in congruent and incongruent conditions) modulates the motor response of participants. Deficits in facial expressions recognition are one of the most common cognitive impairments, and they have been extensively studied in several clinical population, including schizophrenia [Barkhof et al. (2015); Gao et al. (2021)], Autism Spectrum Disorder [Loth et al. (2018)], ADHD [Staff et al. (2022)], Parkinson desease [Argaud et al. (2018); Gray and Tickle-Degnen (2010); Ricciardi et al. (2017)] etc. In this light, methodology and results of the present study are promising for future research aiming to deepen the study of affective communication in patients with face recognition impairments or as clinical assessment to predict the progress of a pathological condition.

Appendix B

Behavioural study 1 - pre kinematic experiment

In order to validate experimental stimuli and ascertain that all conditions were perceived as different, a preliminary behavioural study was carried out before the kinematic experiment on forty-eight healthy participants. Participants were required to carefully observe video-clips of the iCub robot requests. After each stimulus, they were asked to answer the question "*How would you describe the iCub robot?*" by choosing three adjectives from a given list.
Considering the first choice, adjectives selected with a percentage lower than or equal to 6.25% were excluded from the statistical analysis. The remaining adjectives, each one with a relative percentage of choice, were analyzed with a Chi square test. For all the experimental conditions, the p-value resulting from the test was <0.001, thus frequencies of choice of adjectives were significantly different. Additionally, standardized residuals showed that the only adjectives whose frequency was greater than the expected one were: *happy* (5.51>|1.96|) for the congruent condition "gentle action-happy facial expression" (Figure 6.5 A1), *angry* (7.46>|1.96|) for the congruent condition "rude action-angry facial expression" (Figure 6.5 A2), *angry* (9.87>|1.96|) for the incongruent condition "gentle action-angry facial expression (Figure 6.5 B1), *happy* (8.43>|1.96|) for the incongruent condition "rude action-happy facial expression (Figure 6.5 B2)".
These results indicate that, independently from the action vitality forms observed (gentle/rude), the face guided the choice of participants in the description of the iCub robot: the positive and negative facial expressions modified the perception of action vitality forms. The same stimuli were used for the kinematic experiment, hypothesizing that the effect found during perception would be also present during action execution towards the iCub robot.

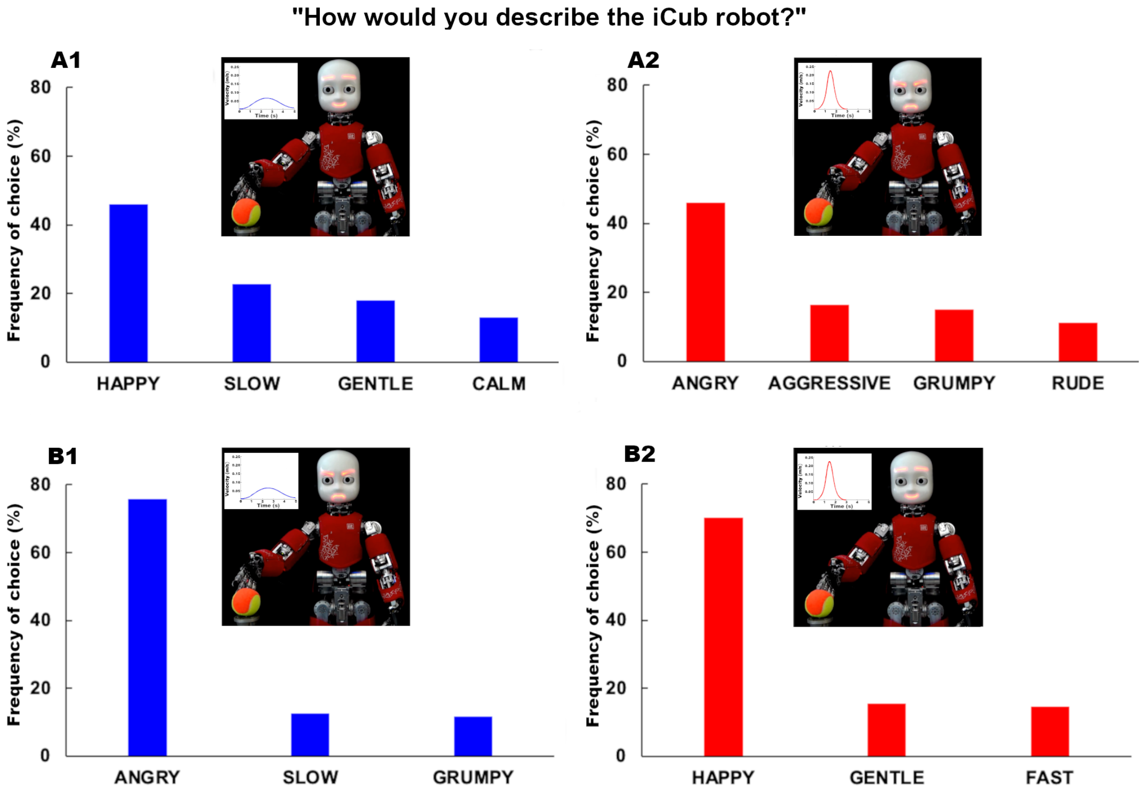


Figure B.1 Results of the first behavioral study. Bars represent frequency of choice of adjectives (%) used to describe each iCub robot request. Congruent conditions: gentle action-happy facial expression (A1), rude action-angry facial expression (A2). Incongruent conditions: gentle action-angry facial expression (B1), rude action-gentle facial expression (B2). Figure adapted from Lombardi et al. (2024).

Behavioural study 2 - post kinematic experiment

In order to deeply understand the influence of the iCub robot request on the affective state of participants, an additional behavioral study was carried out. After the kinematic experiment, each participant was required to observe again all the visual stimuli and answer to the question "*How did you feel after this iCub request?*" by choosing three adjectives from a given list. Considering the first choice, adjectives selected with a percentage lower than or equal to 5% were excluded from the statistical analysis. The remaining adjectives were analyzed with a Chi square test. The only condition whose frequencies of choice of adjectives were significantly different (p=0.04) was the incongruent condition "gentle action-angry facial expression". Standardized residuals showed that the only adjectives whose frequency was greater than the expected one were *calm* (2.44>|1.96|) for the congruent condition "gentle action-happy facial expression" (Figure 6.6 A1), *threatened* (2.11>|1.96|) for the incongruent condition "gentle action-angry facial expression (Figure 6.6 B1) and *surprised* (2.14>|1.96|) for the congruent condition "rude action-happy facial expression (Figure 6.6 B2). For the congruent condition "rude action-angry facial expression" (Figure 6.6 A2), standardized residuals did not show adjectives whose frequency was greater than the expected one.

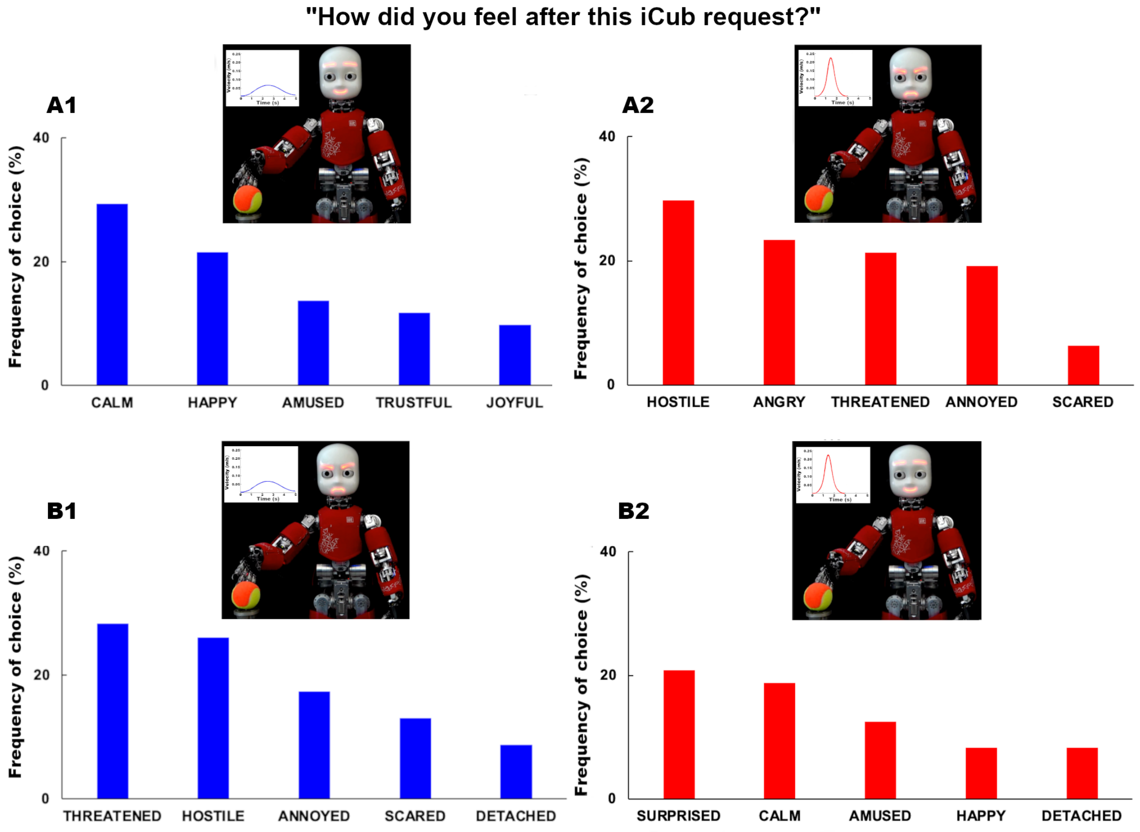


Figure B.2 Results of the second behavioral study. Bars represent frequency of choice of adjectives (%) which participants used to describe their attitude towards the iCub robot after each request. Congruent conditions: gentle action-happy facial expression (A1), rude action-angry facial expression (A2). Incongruent conditions: gentle action-angry facial expression (B1), rude action-gentle facial expression (B2). Figure adapted from Lombardi et al. (2024).

Part IV

Conclusion

Chapter 7

Final discussion

7.1 Overview

Social interactions are characterized by the capacity to communicate our internal states and to evaluate those of others. This behavioural exchange is highly based on actions dynamics of the interactants, which have been defined by Daniel Stern with the term "vitality forms". In addition to the goal (*what*) and intention (*why*) of actions, vitality forms constitute a third fundamental aspect, i.e. *how* actions are performed, reflecting the affective state of the agent. Despite their pivotal role in human communication, vitality forms are still poorly known from researchers and their potential application in different fields needs to be explored.

By emerging at the intersection of neuroscience and robotics, the present thesis involves an interdisciplinary approach to investigate vitality forms from different points of view.

After an extensive introduction to the field of action understanding, particularly emphasizing the role of the mirror mechanism in human action comprehension and expression, Chapter 2 presented the definition of vitality forms and highlighted the preliminary outcomes on the topic. Such research framework served as a reference for the experimental part of my dissertation, outlined from Chapter 3 to Chapter 6.

Chapters 3 and 4 assessed RO1, by deeply investigating the activity of brain areas involved in vitality forms processing in humans. Firstly, I presented the main fMRI study carried out during the first year of my PhD aiming to assess the role of the dorso-central insula and middle cingulate cortex during the observation and execution of actions conveying vitality forms. Secondly, starting from results achieved in this fMRI experiment, Chapter 4 presented a Dynamic Causal Modelling study carried out during my visiting research period at the University College of London to investigate the effective connectivity of areas related to action goal and form processing.

Once achieved a rich comprehension of the brain mechanisms underlying the encoding of vitality forms in humans, the experiment presented in Chapter 5 assessed RO2, by investigating whether and how vitality forms expressed by an agent could influence the action perception (action estimation) and the motor response (action execution) of the receiver during social interactions.

Finally, starting from results achieved in human—human interactions, Chapter 6 proposed the iCub robot as novel tool to achieve RO3, i.e. investigating the influence of vitality forms in human—robot interactions. Particularly, leveraging on its controllability, the iCub robot became an ideal agent to assess how additional social cues such as positive and negative facial expressions could influence the perception of vitality forms conveyed through its gestures and consequently modulating the motor response of the human receiver.

7.2 Achievement of Research Objectives

After restating the primary objective and the approach undertaken in this thesis, I summarize here the main findings and contributions for each research objective presented in the first chapter and further deepened in each experimental chapter (3-6) and corresponding published works.

7.2.1 RO1: Investigate the activity of brain areas involved in vitality forms processing in humans

As first research objective, I deeply investigated the neural mechanism underlying the ability to process *how* actions are performed. Particularly, starting from previous findings, the main aim was to clarify the role of the dorso-central insula and middle cingulate cortex during the observation and execution of actions endowed with vitality forms. Thanks to the collaboration with the Department of Neuroscience of the University of Parma, I used functional magnetic resonance (fMRI) as principal neuroimaging technique to assess this goal. In the main fMRI study carried out in the first year of PhD, participants were asked to perform two main tasks: Observation and Execution. Specifically, taking inspiration from previous fMRI experiments, participants could either observe or execute actions (passing actions) conveying rude or gentle vitality forms (Vitality forms conditions) or observe/execute the same actions without vitality forms (Control conditions). This simple 2x2 experimental design enabled to understand brain areas selective for vitality forms vs control processing. Moreover, the presence of two tasks per condition allowed to assess if the same neural substrates were involved in both observation and execution of action vitality forms and thus characterized by a specific mirror mechanism.

Results indicate that the dorso-central insula (DCI) and the middle cingulate cortex (MCC) are selectively involved in the observation and execution of actions performed with different vitality forms. Furthermore, thanks to a multifiber tractography investigation, these two brain areas have been proven to be anatomically connected. The most important result comes from a voxel-based analysis, showing that a large proportion of the most active voxels ("high correlated mirror voxels") in both DCI and MCC are similarly activated during the observation and execution tasks, indicating the presence of mirror properties in this brain circuit. While the general contrasts Vitality Observation vs Control Observation (VF OBS vs CT OBS) and Vitality Execution vs Control execution (VF EXE vs CT EXE) reveal the activation of these brain areas and the classical parieto-frontal network, the conjunction analysis carried out to highlight selective activations for vitality forms processing shows a specific activity of DCI and MCC only. This means that during action observation and execution two main circuits with mirror properties are active, one related to the understanding of the action goal and the second one related to the encoding of specific vitality forms characterizing the action.

In this view and based on the rich literature on anatomical connections of the insula and the parieto-frontal network, RO1 was further deepened during my visiting research period at the Wellcome Centre for Human Neuroimaging of the University College of London by investigating the effective connectivity, i.e. causal influence, of these two brain circuits during actions processing with Dynamic Causal Modelling. Results of this DCM study show that, during action observation, pSTS (i.e.visual) area gives origin to two visual streams, one towards IPL, for action goal, and one towards DCI, for action vitality forms. During action execution, motor information starts from PM and also gives origin to two streams: one, related to the action goal, towards IPL and one, concerning the action vitality form, towards DCI. It is important to note that the significant modulation of DCI was absent during the observation and execution of control actions, highlighting its specificity for vitality forms processing only. These data clearly indicate that, during the observation and execution of goal directed actions, two streams arise from visual and motor areas: a dorsal stream encoding the action goal and a ventral stream encoding the action vitality form.

7.2.2 RO2: Investigate the influence of vitality forms on action perception and expression during human—human interactions

As demonstrated in the first experimental part of my dissertation, action perception and execution rely on the same neural correlates and thanks to the mirror mechanism representations of actions (action goal and action form) performed by others are transformed into one own motor representation. In this regard, as second objective I investigated whether and how this mechanism may affect the behaviour of human participants during social interactions. Particularly, RO2 proposed to quantify the effect of vitality forms expressed by an agent, through different modalities, on both action perception and action execution of a human receiver. To this purpose, the experiment reproduced an interactive scenario in which participants were stimulated with requests made through a physical contact or vocally and conveying rude or gentle vitality forms, and then they were asked to estimate the end of a passing action observed in a monitor (action estimation task) or to perform an action in front of it (action execution task) with the intention to pass an object to the other person presented in the video. Results of the action estimation task indicate that the perception of a gentle request increases the duration of a rude action subsequently observed, while the perception of a rude request decreases the duration of the same action performed gently. Additionally, during the action execution task, accordingly with the perceived vitality form, participants modulate their motor response. Particularly, after a rude request, their passing action is characterized by a higher velocity peak and a wider trajectory. Conversely, after a gentle request, the same action is performed with a lower velocity peak and covers a smaller distance. On the basis of these findings, it can be hypothesized that the neural processing of vitality forms has an important effect on our everyday life interactions: it allows to understand vitality forms expressed by others, remapping them in our motor schema and influencing their internal representation. Such mechanism would allow to prepare adequate responses and obtain a fluidity of interaction with other individuals.

7.2.3 RO3: Investigate the influence of vitality forms on action perception and expression during human—robot interactions

Once ascertained the effect of vitality forms in human—human interactions, from both a cognitive and behavioural point of view, an interesting point consisted in the evaluation of the same influence during human—robot interactions. Starting from findings previously obtained by our research group [Vannucci et al. (2018)], RO3 proposed to deepen the study of vitality forms in human—robot interaction by investigating the effect of additional social cues such as facial expressions, which have been always excluded in both neuroimaging and behavioural studies concerning vitality forms. Specifically, we asked if the influence of the iCub robot requests (gentle/rude actions) on the motor response of human participants could be modified if the robot displayed also positive or negative facial expressions. The humanoid robot iCub represents an ideal agent for the investigation of this effect. Indeed, thanks to its controllability and consequent repeatability, necessary in quantitative investigation, four conditions were generated and presented to participants: two congruent conditions in which the iCub request was conveyed through a gentle action with a happy facial expression (positive valence) or a rude action with an angry facial expression (negative valence); two incongruent conditions in which the action vitality form and the facial expression displayed by the iCub robot were characterized by opposite valence. These four requests were tested in a kinematic experiment, to verify the modulation of participants kinematic features, and in two behavioural studies, to evaluate how participants perceived each request and their feelings towards the iCub robot.

Results show that the iCub facial expressions significantly modulate participants motor response. Particularly, the observation of a happy facial expression, associated to a rude action, decreases specific kinematic parameters such as velocity, acceleration and maximum height of movement. In contrast, the observation of an angry facial expression, associated to a gentle action, increases the same kinematic parameters. Also the reaction time with which participants initiate the motor response towards the iCub robot is significantly modulated by both vitality forms and facial expressions. These data indicate that positive and negative facial expressions modify the perception of action vitality forms, by giving an additional social context to the interaction. Also results of the behavioural studies confirm this suggestion, showing an enhancement of action vitality forms effect during congruent conditions, and a modification of vitality forms perception during incongruent conditions. For example, a gentle action, characterized by low velocity and acceleration peaks and always attributed to a positive attitude of the agent, was perceived as more gentle when associated to a happy facial expression but changed the perception of the attitude communicated by the agent (e.g. hostile), when associated to an angry facial expression.

This final experiment remarks the important role of additional information such as facial expressions in the understanding other’s internal states and their significant modulation of human behaviour when combined with different action vitality forms. More interestingly, it demonstrated the effective use of the iCub robot as agent of the interaction, proposing it as ideal probe for future investigations on the topic.

7.3 Final considerations, limitations, and future works

The work presented in this thesis aims to contribute to the knowledge of vitality forms, highlighting their key role in the understanding of others' internal states during social interactions. Results obtained during my PhD demonstrate the existence of an affective contagion, through which vitality forms expressed by an agent influence action perception and execution of human individuals, activating specific neural circuits in their brain. Experiments carried out during my PhD represent one of the first attempts to demonstrate the importance of the action vitality form in the study of human behaviour and its promising application for the achievement of effective/affective communication between humans and robots.
During the three years of my PhD, the collaboration between the Italian Institute of Technology of Genova and the Department of Neuroscience of the University of Parma allowed me to follow an interdisciplinary approach for the investigation of my research topic, enabling to build a robust connection between different fields, such as cognitive neuroscience and robotics.
I concluded each chapter of the dissertation by presenting some limitations specific for each experiment. However, other more general issues must be here discussed and tackled in the future. First, in all the neuroimaging, kinematic and behavioural studies presented, only two vitality forms were considered, named gentle and rude. However, each action can be characterized by different levels of gentleness and rudeness, depending on the even minimal variation of particular kinematic features. Future research could further expand and deepen this aspect, by enabling to explore the influence of a wider dimension of vitality forms, such as hesitant, vigorous, hostile, furious etc. Moreover, the present experiments investigated the influence of observation and execution of basic goal directed actions, such as passing objects. In this view, together with a richer repertoire of vitality forms, also the processing of additional actions may be further explored in future experiments.
Additionally, the possibility to endow robots with the ability to express vitality forms introduces several ethical issues that should be discussed. With a minimal of social cues users tend to personify socially interactive robots, which fosters humans to form unidirectional emotional bonds with them [de Graaf (2016)]. Some researchers even claim that the ability of robots to respond with human-like social cues should already be regarded as a form of deception [Wallach and Allen (2008)]. It can be argued that robot deception might be legitimate under some circumstances, for instance, when the goal is to make the human partner feel positive [Arkin et al. (2012)]. Other researchers agree that robot deception is ethically problematic, no matter what [Sharkey and Sharkey (2012a)]. Isolation is seen as another

main reason for resisting the use of social robots especially among older people [Körtner (2016); Parks (2010)]. For example, while there is evidence that human companionship help to delay the onset of dementia, evidence with social robots is lacking [Sharkey and Sharkey (2012b)]. Robotics researchers and ethicists have also raised questions about the potential influence of social robots on children's development. Since children have a strong social drive, there is a risk that they may overestimate a robot's abilities, causing the child to spend too much time with robots at the expense of interactions with human beings, developing strong psychological bonds and emotional attachment [Langer et al. (2023); Riek and Howard (2014)].

To conclude, research and methodology proposed in this PhD thesis lay solid bases for future studies. The presence of a specific circuit for vitality forms encoding (RO1) and consequent modulation of action perception and execution during social interactions (RO2) open new questions for clinical research. Particularly, it could be interesting to explore if the same significant modulation of DCI during vitality forms observation/execution may be absent in the presence of social impairments, such as Autism Spectrum Disorder, Schizophrenia, Borderline Personality Disorder etc., with a subsequent deficit in the expression of vitality forms and their perception during interactions with others. Furthermore, in the next few years humans will increasingly interact with humanoid robots in several scenarios, such as rehabilitation. In this view the interdisciplinary approach used in this thesis, based on a solid collaboration between neuroscientists and roboticists, will be crucial to enhance the implementation – and ultimately effectiveness – of social robots [Langer and Levy-Tzedek (2021)]. Taking into consideration ethical issues discussed above, a fascinating possibility is that, besides vitality forms expression, new generations of humanoids will be endowed also with the capacity to comprehend vitality forms expressed by humans interacting with them more effectively.

7.4 Publications and Dissemination

The publications made during my PhD period are listed below.

Published

- **Lombardi G.**, Sciutti A., Rea F., Vannucci F., Di Cesare G. Humanoid facial expressions as a tool to study human behaviour. Scientific Reports, 2024; 14(133). https://doi.org/10.1038/s41598-023-45825-6.
- **Lombardi G.***, Gerbella M.*, Marchi M., Sciutti A., Rizzolatti G., Di Cesare G. Investigating form and content of emotional and non-emotional laughing. Cerebral Cortex, 2022; 33(7), 4164:4172, https://doi.org/10.1093/cercor/bhac334.
- **Lombardi G.**, Zenzeri J., Belgiovine G., Vannucci F., Rea F., Sciutti A., Di Cesare G. The influence of vitality forms on action perception and motor response. Scientific Reports, 2021; 11(1):22576. doi: 10.1038/s41598-021-01924-w.
- Di Cesare G., Marchi M., **Lombardi G.**, Gerbella M., Sciutti, A., Rizzolatti G. The middle cingulate cortex and dorso-central insula: a mirror circuit encoding observation and execution of vitality forms. Proceedings of the National Academic Science of the United States of America, 2021; 118(44):e2111358118. doi: 10.1073/pnas.2111358118.
- Di Cesare G., Pelosi A., Aresta S., **Lombardi G.**, Sciutti, A. Affective contagion: how attitudes expressed by others influence our perception of actions. Frontiers in Human Neuroscience, 2021; 15:712550. doi: 10.3389/fnhum.2021.712550.
- **Lombardi G.**, Sauer M., Di Cesare G. An Affective Perception: How “Vitality Forms” Influence Our Mood. Art Style, Art & Culture International Magazine, 2023; 11(3), 127-139. https://zenodo.org/records/7651433
- **Lombardi G.**, Di Cesare G. From Neuroscience to Art. The Role of “Vitality Forms” in the investigation of multimodality. Art Style, Art & Culture International Magazine, 2022; 10(10), 11-23. https://zenodo.org/records/7020465

Submitted / in preparation

- Di Cesare G., **Lombardi G.***, Zeidman P., Urgen B.A., Sciutti A., Friston K., Rizzolatti G. Two distinct networks for encoding goals and forms of action: an effective connectivity study.
- Vannucci F.*, **Lombardi G.***, Rea F., Sandini G., Di Cesare G., Sciutti A. Humanoid attitudes influence humans in video and live interactions.

Conferences and events in which I presented my research are listed below:

- **Poster presentation** at the 19th National Congress of the Italian Society for Neuroscience (SINS), Online 9-11 September 2021. Poster title: Evidence of mirror mechanism for vitality forms encoding in the insula and cingulate cortices.
- **Symposium speaker** at the Italian Association of Cognitive Science (AISC) midterm conference, Parma (Italy), 22-24 June 2022. Talk title: Towards an affective human robot interaction: the role of vitality forms.
- **Symposium speaker** at the XXX National Congress SIPF, Udine (Italy), 15-17 September 2022. Talk title: The influence of vitality forms on action perception and motor response.
- **Poster presentation** at XXIX AIP Congress, Lucca (Italy), 18-20 September 2023. Poster title: Humanoid facial expressions: a tool to study human behaviour

References

Adolphs, R. (2002). Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behav Cogn Neurosci Rev*, 1(1):21–62.

Aglioti, S. M., Cesari, P., Romani, M., and Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nat Neurosci*, 11(9):1109–16.

Almashaikhi, T., Rheims, S., Jung, J., Ostrowsky-Coste, K., Montavont, A., De Bellescize, J., Arzimanoglou, A., Keo Kosal, P., Guénot, M., Bertrand, O., and Ryvlin, P. (2014). Functional connectivity of insular efferences. *Hum Brain Mapp.*, 35(10):5279–94.

Ansuini, C., Santello, M., Massaccesi, S., and Castiello, U. (2006). Effects of end-goal on hand shaping. *J Neurophysiol.*, 95(4):2456–65.

Argaud, S., Vérin, M., Sauleau, P., and Grandjean, D. (2018). Facial emotion recognition in parkinson's disease: A review and new hypotheses. *Mov Disord.*, 33(4):554–567.

Arkin, R. C., Ulam, P., and Wagner, A. R. (2012). Moral decision making in auton- omous systems: Enforcement, moral emotions, dignity, trust, and deception. *Proc IEEE*, 100:571–89.

Barkhof, E., de Sonneville, L. M. J., Meijer, C. J., and de Haan, L. (2015). Specificity of facial emotion recognition impairments in patients with multi-episode schizophrenia. *Schizophr Res Cogn.*, 2(1):12–19.

Bastos, R. N., Penate, X., Bates, M., Hammond, D., and Barr, F. A. (2012). Cyk4 inhibits rac1-dependent pak1 and arhgef7 effector pathways during cytokinesis. *J Cell Biol.*, 198(5):865–80.

Becchio, C., Sartori, L., Bulgheroni, M., and Castiello, U. (2008). The case of dr. jekyll and mr. hyde: a kinematic study on social intention. *Conscious Cogn.*, 17(3):557–64.

Bernstein, M. and Yovel, G. (2015). Two neural pathways of face processing: a crit- ical evaluation of current models. *Neurosci Biobehav Rev.*, 55:536–546.

Berstein, N. A. (1996). *Dexterity and its Development*. Psychology Press, Taylor and Francis Group.

Bonini, L., Rozzi, S., Serventi, F. U., Simone, L., Ferrari, P. F., and Fogassi, L. (2010). Ventral premotor and inferior parietal cortices make distinct contribution to action organization and intention understanding. *Cereb Cortex.*, 20(6):1372–85.

Bonini, L., Serventi, F. U., Rozzi, S., Ferrari, P. F., and Fogassi, L. (2011). Grasping neurons of monkey parietal and premotor cortices encode action goals at distinct levels of abstraction during complex action sequences. *J Neurosci.*, 31(15):5876–86.

Bradley, R. M. and Mistretta, C. M. (1975). Fetal sensory receptors. *Physiological Reviews*, 55(3):352–382.

Bruce, V. and Young, A. (1976). Understanding face recognition. *Brit J Psychol.*, 77(3):305–327.

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., and Haggard, P. (2005). Action observation and acquired motor skills: an fmri study with expert dancers. *Cereb Cortex.*, 15(8):1243–9.

Calvo-Merino, B., Grèzes, J., Glaser, D. E., , Passingham, R. E., and Haggard, P. (2006). Seeing or doing? influence of visual and motor familiarity in action observation. *Curr Biol.*, 16(19):1905–10.

Cappella, J. N. and Planalp, S. (1981). Talk and silence sequences in informal conversations: Iii. interspeaker influence. *Human Communication Research*, 7(2):117–132.

Caruana, F., Gerbella, M., Avanzini, P., Gozzo, F., Pelliccia, V., Mai, R., Abdollahi, R. O., Cardinale, F., Sartori, I., Lo Russo, G., and Rizzolatti, G. (2018). Motor and emotional behaviours elicited by electrical stimulation of the human cingulate cortex. *Brain.*, 141(10):3035–3051.

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.

Casile, A., Dayan, E., Caggiano, V., Hendler, T., Flash, T., and Giese, M. A. (2007). Neuronal encoding of human kinematic invariants during action observation. *Cereb Cortex.*, 20(7):1647–55.

Caspers, S., Zilles, K., Laird, A. R., and Eickhoff, S. B. (2010). Ale meta-analysis of action observation and imitation in the human brain. *Neuroimage*, 50(3):1148–67.

Cohen, R. G. and Rosenbaum, D. A. (2004). Where objects are grasped reveals how grasps are planned: Generation and recall of motor plans. *Experimental Brain Research*, 157:486–495.

Craig, A. D. (2002). How do you feel? interoception: the sense of the physiological condition of the body. *Nat Rev Neurosci.*, 3(8):655–66.

Cross, E. S., Kraemer, D. J., Hamilton, A. F., Kelley, W. M., and Grafton, S. T. (2009). Sensitivity of the action observation network to physical and observational learning. *Cerebral Cortex*, 19(2):315–26.

Darwin, C. (1872). *The Expression of the Emotions in Man and Animals.* John Murray, London.

de Gelder, B., Snyder, J., Greve, D., Gerard, G., and Hadjikhani, N. (2004). Fear fosters flight: a mechanism for fear contagion when perceiving emotion expressed by a whole body. *Proc Natl Acad Sci U S A.*, 101(47):16701–6.

de Graaf, M. M. A. (2016). An ethical evaluation of human–robot relationships. *Int J of Soc Robotics*, 8:589–598.

Di Cesare, G., De Stefani, E., Gentilucci, M., and De Marco, D. (2017). Vitality forms expressed by others modulate our own motor response: A kinematic study. *Front Hum Neurosci.*, 11:565.

Di Cesare, G., Di Dio, C., Marchi, M., and Rizzolatti, G. (2015). Expressing our internal states and understanding those of others. *Proc Natl Acad Sci U S A.*, 112(33):10331–5.

Di Cesare, G., Di Dio, C., Rochat, M. J., Sinigaglia, C., Bruschweiler-Stern, N., Stern, D. N., and Rizzolatti, G. (2013). The neural correlates of 'vitality form' recognition: an fmri study: this work is dedicated to daniel stern, whose immeasurable contribution to science has inspired our research. *Soc Cogn Affect Neurosci.*, 9(7):951–60.

Di Cesare, G., Gerbella, M., and Rizzolatti, G. (2020a). The neural bases of vitality forms. *Natl Sci Rev.*, 7(1):202–213.

Di Cesare, G., Marchi, M., Lombardi, G., Gerbella, M., Sciutti, A., and Rizzolatti, G. (2021a). The middle cingulate cortex and dorso-central insula: A mirror circuit encoding observation and execution of vitality forms. *Proc Natl Acad Sci U S A.*, 118(44):e2111358118.

Di Cesare, G., Marchi, M., Pinardi, C., and Rizzolatti, G. (2019a). Understanding the attitude of others by hearing action sounds: the role of the insula. *Sci Rep*, 9(1):14430.

Di Cesare, G., Pelosi, A., Aresta, S. M., Lombardi, G., and Sciutti, A. (2021b). Affective contagion: How attitudes expressed by others influence our perception of actions. *Front Hum Neurosci.*, 15:712550.

Di Cesare, G., Pinardi, C., Carapelli, C., Caruana, F., Marchi, M., Gerbella, M., and Rizzolatti, G. (2019b). Insula connections with the parieto-frontal circuit for generating arm actions in humans and macaque monkeys. *Cereb Cortex.*, 29(5):2140–2147.

Di Cesare, G., Vannucci, F., Rea, F., Sciutti, A., and Sandini, G. (2020b). How attitudes generated by humanoid robots shape human brain activity. *Sci Rep.*, 10(1):16928.

Di Dio, C., Di Cesare, G., Higuchi, S., Roberts, N., Vogt, S., and Rizzolatti, G. (2013). The neural correlates of velocity processing during the observation of a biological effector in the parietal and premotor cortex. *Neuroimage.*, 64:425–36.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., and Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Exp. Brain Res.*, 91:176–180.

Dimberg, U. (1982). Facial reactions to facial expressions. *Psychophysiology.*, 19(6):643–647.

Dimberg, U. and Thunberg, M. (2012). Empathy, emotional contagion, and rapid facial reactions to angry and happy facial expressions. *Psych J.*, 1(2):118–127.

Dimberg, U., Thunberg, M., and Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychol Sci.*, 11(1):86–89.

Duff, E. P., Cunnington, R., and Egan, G. F. (2007). Rex: response exploration for neuroimaging datasets. *Neuroinformatics.*, 5(4):223–234.

Efron, D. (1941). *Gesture and environment.* Oxford, England: King's Crown Press.

Eizicovits, D., Edan, Y., Tabak, I., and Levy-Tzedek, S. (2018). Robotic gaming prototype for upper limb exercise: Effects of age and embodiment on user preferences and movement. *Restor Neurol Neurosci.*, 36(2):261–274.

Ekman, P. (1993). Facial expression and emotion. *Am Psychol.*, 48(4):384–92.

Ekman, P., Friesen, W. V., and Ellsworth, P. C. (1972). *Emotion in the human face: Guidelines for research and an integration of findings.* New York: Pergamon.

Ferrari, P. F. and Rizzolatti, G. (2015). *New Frontiers in Mirror Neurons Research.* Oxford Academic.

Ferrari, P. F., Rozzi, S., and Fogassi, L. (2005). Mirror neurons responding to observation of actions made with tools in monkey ventral premotor cortex. *J Cogn Neurosci.*, 17(2):212–226.

Field, T. (2010). Touch for socioemotional and physical well-being: A review. *Developmental Review*, 30(4):367–383.

Filimon, F., Nelson, J. D., Hagler, D. J., and Sereno, M. I. (2007). Human cortical representations for reaching: mirror neurons for execution, observation, and imagery. *Neuroimage.*, 37(4):1315–28.

Fischer, T., Puigbò, J. Y., Camilleri, D., Nguyen, P. D. H., Moulin-Frier, C., Lallée, S., Metta, G., Prescott, T. J., Demiris, Y., and Verschure, P. F. M. J. (2018). icub-hri: A software framework for complex human-robot interaction scenarios on the icub humanoid robot. *Front Robot AI.*, 5:22.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., and Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science.*, 308(5722):662–7.

Friesen, W., Ekman, P., and Wallbott, H. (1979). Measuring hand movements. *Journal of Nonverbal Behavior*, 4:97–112.

Fujiwara, K. and Daibo, I. (2016). Evaluating interpersonal synchrony: Wavelet transform toward an unstructured conversation. *Front Psychol.*, 7:516.

Furl, N., Hadj-Bouziane, F., Liu, N., Averbeck, B. B., and Ungerleider, L. G. (2012). Dynamic and static facial expressions decoded from motion- sensitive areas in the macaque monkey. *J Neurosci.*, 32(45):15952–15962.

Gallace, A. and Spence, C. (2010). The science of interpersonal touch: an overview. *Neurosci. Biobehav. Rev.*, 34:246–59.

Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119:593–609.

Gao, Z., Zhao, W., Liu, S., Liu, Z., Yang, C., and Xu, Y. (2021). Facial emotion recognition in schizophrenia. *Front Psychiatry.*, 12(1):633717.

Gazzola, V., Aziz-Zadeh, L., and Keysers, C. (2006). Empathy and the somatotopic auditory mirror system in humans. *Current Biology*, 16(18):14430.

Gentilucci, M., Fogassi, L., Luppino, G., Matelli, M., Camarda, R., and Rizzolatti, G. (1988). Functional organization of inferior area 6 in the macaque monkey. i. somatotopy and the control of proximal movements. *Exp Brain Res.*, 71(3):475–490.

Gentilucci, M., Negrotti, A., and Gangitano, M. (1997). Planning an action. *Exp Brain Res.*, 115(1):116–28.

Gielniak, M. J., Liu, C. K., and Thomaz, A. L. (2013). Generating human-like motion for robots. *Int. J. Rob. Res.*

Gray, H. M. and Tickle-Degnen, L. (2010). A meta-analysis of performance on emotion recognition tasks in parkinson's disease. *Neuropsychology*, 24(2):176–91.

Grosbras, M. H., Beaton, S., and Eickhoff, S. B. (2012). Brain regions involved in human movement perception: a quantitative voxel-based meta-analysis. *Hum Brain Mapp.*, 33(2):431–54.

Hatfield, E., Cacioppo, J. T., and Rapson, R. L. (1993). Emotional contagion. *Curr Dir Psychol Sci.*, 2(3):96–100.

Hertenstein, M. J. (2002). Touch: Its communicative functions in infancy. *Human Development*, 45(2):70–94.

Hertenstein, M. J., Holmes, R., McCullough, M., and Keltner, D. (2009). The communication of emotion via touch. *Emotion*, 9:566–73.

Hertenstein, M. J., Keltner, D., App, B., Bulleit, B. A., and Jaskolka, A. R. (2002). Touch communicates distinct emotions. *Human Development*, 6:528–33.

Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. (1999). Bayesian model averaging: a tutorial (with comments by m. clyde, david draper and ei george, and a rejoinder by the authors. *Statistical science*, 14(4):382–417.

Jezzini, A., Caruana, F., Stoianov, I., Gallese, V., and Rizzolatti, G. (2012). Functional organization of the insula and inner perisylvian regions. *Proc Natl Acad Sci U S A.*, 109(25):10077–82.

Jezzini, A., Rozzi, S., Borra, E., Gallese, V., Caruana, F., and Gerbella, M. (2015). A shared neural network for emotional expression and perception: an anatomical study in the macaque monkey. *Front Behav Neurosci.*, 9:243.

Jones, A. L., Kramer, R. S., and Ward, R. (2012). Signals of personality and health: the contributions of facial shape, skin texture, and viewing angle. *J Exp Psychol Hum Percept Perform.*, 38(6):1353–61.

Karg, M., Samadani, A. A., Gorbet, R., Kuhnlenz, K., Hoey, J., and Kulic, D. (2013). Body movements for affective expression: A survey of automatic recognition and generation. *IEEE Trans. Affect. Comput.*

Keysers, C. and Gazzola, V. (2009). Expanding the mirror: vicarious activity for action, emotion and sensation. *Curr. Opin. Neurobiol.*, 19(6):666–671.

Kirsch, L. P. and Cross, E. S. (2015). Additive routes to action learning: Layering experience shapes engagement of the action observation network. *Cerebral Cortex*, 25(12):4799–811.

Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., and Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, 297(5582):846–848.

Kurata, K. and Tanji, J. (1986). Premotor cortex neurons in macaques: activity before distal and proximal forelimb movements. *Journal of Neuroscience*, 6(2):403–411.

Kurth, F., Zilles, K., Fox, P. T., Laird, A. R., and Eickhoff, S. B. (2010). A link between the systems: functional differentiation and integration within the human insula revealed by meta-analysis. *Brain Struct Funct.*, 214(5-6):519–34.

Körtner, T. (2016). Ethical challenges in the use of social service robots for elderly people. *Z Gerontol Geriat*, 49:303–307.

Langer, A. and Levy-Tzedek, S. (2020). Priming and timing in human-robot interactions. *Modelling Human Motion: From Human Perception to Robot Design, Springer*, pages 335–350.

Langer, A. and Levy-Tzedek, S. (2021). Emerging roles for social robots in rehabilitation: Current directions. *ACM Transactions on Human-Robot Interaction*, 10(4):1–4.

Langer, A., Marshall, P. J., and Levy-Tzedek, S. (2023). Ethical considerations in child-robot interactions. *Neurosci Biobehav Rev.*, 151:15230.

Lazzeri, N., Mazzei, D., Ben Moussa, M., Magnenat-Thalmann, N., and De Rossi, D. (2018). The influence of dynamics and speech on understanding humanoid facial expressions. *International Journal of Advanced Robotic Systems.*, 15(4).

Leppanen, J. M. and Hietanen, J. K. (2004). Positive facial expressions are recognized faster than negative facial expressions, but why? *Psychol Res*, 69:22–29.

Lombardi, G., Gerbella, M., Marchi, M., Sciutti, A., Rizzolatti, G., and Di Cesare, G. (2017). Investigating form and content of emotional and non-emotional laughing. *Cerebral Cortex*, 33(7):4164–4172.

Lombardi, G., Sciutti, A., Rea, F., Vannucci, F., and Di Cesare, G. (2024). Humanoid facial expressions as a tool to study human behaviour. *Sci Rep*, 14(133).

Lombardi, G., Zenzeri, J., Belgiovine, G., Vannucci, F., Rea, F., Sciutti, A., and Di Cesare, G. (2021). The influence of vitality forms on action perception and motor response. *Sci Rep*, 11(1):22576.

Loth, E., Garrido, L., Ahmad, J., Watson, E., Duff, A., and Duchaine, B. (2018). Facial expression recognition as a candidate marker for autism spectrum disorder: how frequent and severe are deficits? *Mol Autism.*, 9:7.

Madhavan, S. and Stoykov, M. E. (2017). Editorial: Motor priming for motor recovery: Neural mechanisms and clinical perspectives. *Front Neurol.*, 8:448.

Maranesi, M., Rodà, F., Bonini, L., Rozzi, S., Ferrari, P. F., Fogassi, L., and Coudè, G. (2012). Anatomo-functional organization of the ventral primary motor and premotor cortex in the macaque monkey. *Eur J Neurosci.*, 36(10):3376–87.

Marteniuk, R. G., MacKenzie, C. L., Jeannerod, M., Athenes, S., and Dugas, C. (1987). Constraints on human arm movement trajectories. *Can J Psychol.*, 41(3):365–78.

Marx, V. and Nagy, E. (2015). Fetal behavioural responses to maternal voice and touch. *PLoS One*, 10(6).

Matsumoto, D. (2006). *Culture and nonverbal behavior.* The SAGE handbook of nonverbal communication.

Mesulam, M. M. and Mufson, E. J. (1982). Insula of the old world monkey. iii: Efferent cortical output and comments on function. *J Comp Neurol.*, 212(1):38–52.

Metta, G., Natale, L., Nori, F., Sandini, G., Vernon, D., Fadiga, L., von Hofsten, C., Rosander, K., Lopes, M., Santos-Victor, J., Bernardino, A., and Montesano, L. (2010). The icub humanoid robot: an open-systems platform for research in cognitive development. *Neural Netw.*, 23(8-9):1125–34.

Molenberghs, P., Cunnington, R., and Mattingley, J. B. (2012). Brain regions with mirror properties: a meta-analysis of 125 human fmri studies. *Neurosci Biobehav Rev*, 36(1):341–349.

Moody, E. J., McIntosh, D. N., Mann, L. J., and Weisser, K. R. (2007). More than mere mimicry? the influence of emotion on rapid facial reactions to faces. *Emotion.*, 7(2):447–57.

Naish, K. R., Reader, A. T., Houston-Price, C., Bremner, A. J., and Holmes, N. P. (2013). To eat or not to eat? kinematics and muscle activity of reach-to-grasp movements are influenced by the action goal, but observers do not detect these differences. *Exp Brain Res.*, 225(2):261–75.

Nelissen, K., Borra, E., Gerbella, M., Rozzi, S., Luppino, G., Vanduffel, W., Rizzolatti, G., and Orban, G. A. (2011). Action observation circuits in the macaque monkey cortex. *J Neurosci.*, 31(10):3743–56.

Ou, Y., Dai, P., Zhou, X., Tong, X., Yang, L., Zailiang, C., and Beiji, Z. (2022). A strategy of model space search for dynamic causal modeling in task fmri data exploratory analysis. *Phys Eng Sci Med*, 45:867–882.

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

Pani, P., Theys, T., Romero, M. C., and Janssen, P. (2014). Grasping execution and grasping observation activity of single neurons in the macaque anterior intraparietal area. *J Cogn Neurosci.*, 26(10):2342–55.

Park, J. W., Lee, H. S., and Chung, M. J. (2015). Generation of realistic robot facial expressions for human robot interaction. *J Intell Robot Syst*, 78:443–462.

Parks, J. A. (2010). Lifting the burden of women's care work: Should robots replace the "human touch"? *Hypatia*, 25:100–120.

Penny, W., Flandin, G., and Trujillo-Barreto, N. (2007a). Bayesian comparison of spatially regularised general linear models. *Hum Brain Mapp.*, 28(4):275–93.

Penny, W. D., Friston, K. J., Ashburne, J., Kiebel, S. J., and Nichols, T. E. (2007b). *Statistical Parametric Mapping: The Analysis of Functional Brain Images*.

Pinilla, A., Tamayo, R. M., and Neira, J. (2020). How do induced affective states bias emotional contagion to faces? a three-dimensional model. *Front Psychol.*, 11:97.

Quesque, F. and Coello, Y. (2014). For your eyes only: effect of confederate's eye level on reach-to-grasp action. *Front Psychol.*, 5:1407.

Quesque, F., Mignon, A., and Coello, Y. (2017). Cooperative and competitive contexts do not modify the effect of social intention on motor action. *Conscious Cogn.*, 56:91–99.

Ricciardi, L., Visco-Comandini, F., Erro, R., Morgante, F., Bologna, M., Fasano, A., Ricciardi, D., Edwards, M. J., and Kilner, J. (2017). Facial emotion recognition and expression in parkinson's disease: An emotional mirror mechanism? *PLoS One.*, 12(1):554–567.

Riek, L. D. and Howard, D. (2014). A code of ethics for the human-robot interaction profession. *We Robot Conference.*

Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., and Matelli, M. (1988). Functional organization of inferior area 6 in the macaque monkey. ii. area f5 and the control of distal movements. *Exp Brain Res.*, 71(3):491–507.

Rizzolatti, G., Cattaneo, L., Fabbri-Destro, M., and Rozzi, S. (2014). Cortical mechanisms underlying the organization of goal-directed actions and mirror neuron-based action understanding. *Physiol. Rev.*, 94(2):655–706.

Rizzolatti, G., D'Alessio, A., Marchi, M., and Di Cesare, G. (2021). The neural bases of tactile vitality forms and their modulation by social context. *Sci Rep.*, 11(1):9095.

Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cogn. Brain Res.*, 3:131–141.

Rizzolatti, G. and Fogassi, L. (2014). The mirror mechanism: recent findings and perspectives. *Philos Trans R Soc Lond B Biol Sci.*, 369(1644):20130420.

Rizzolatti, G., Fogassi, L., and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat Rev Neurosci.*, 2(9):661–670.

Rizzolatti, G. and Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nat Rev Neurosci.*, 11(4):264–74.

Rizzolatti, G. and Sinigaglia, C. (2016). The mirror mechanism: a basic principle of brain function. *Nat Rev Neurosci*, 17(12):757–765.

Rochat, M. J., Caruana, F., Jezzini, A., Escola, L., Intskirveli, I., Grammont, F., Gallese, V., Rizzolatti, G., and Umiltà, M. A. (2010). Responses of mirror neurons in area f5 to hand and tool grasping observation. *Exp Brain Res.*, 204(4):605–616.

Rosenbaum, D. A., Chapman, K. M., Weigelt, M., Weiss, D. J., and van der Wel, R. (2012). Cognition, action, and object manipulation. *Psychol Bull.*, 138(5):924–46.

Rosenbaum, D. A., Halloran, E. S., and Cohen, R. G. (2006). Grasping movement plans. *Psychonomic Bulletin Review*, 13:918–922.

Rozzi, S., Calzavara, R., Belmalih, A., Borra, E., Gregoriou, G. G., Matelli, M., and Luppino, G. (2006). Cortical connections of the inferior parietal cortical convexity of the macaque monkey. *Cereb Cortex.*, 16(10):1389–417.

Rozzi, S., Ferrari, P. F., Bonini, L., Rizzolatti, G., and Fogassi, L. (2008). Functional organization of inferior parietal lobule convexity in the macaque monkey: electrophysiological characterization of motor, sensory and mirror responses and their correlation with cytoarchitectonic areas. *Eur J Neurosci.*, 28(8):1569–88.

Sandini, G., Sciutti, A., and Rea, F. (2017). Movement-based communication for humanoid-human interaction. *Humanoid Robotics: A Reference*, pages 1–19.

Schmidt, R. C., Fitzpatrick, P., Caron, R., and Mergeche, J. (2011). Understanding social motor coordination. *Hum Mov Sci.*, 30(5):834–845.

Sciutti, A., Mara, M., Tagliasco, V., and Sandini, G. (2018). Humanizing human-robot interac- tion: On the importance of mutual understanding. *IEEE Technol. Soc. Mag.*, 37:22–29.

Sharkey, A. and Sharkey, N. (2012a). Granny and the robots: Ethical issues in robot care for the elderly. *Ethics Inf Technol.*, 14:27–40.

Sharkey, N. and Sharkey, A. (2012b). The eldercare factory. *Gerontology*, 58(3):282–288.

Staff, A. I., Luman, M., van der Oord, S., Bergwerff, C. E., van den Hoofdakker, B. J., and Oosterlaan, J. (2022). Facial emotion recognition impairment predicts social and emotional problems in children with (subthreshold) adhd. *Eur Child Adolesc Psychiatry.*, 31(5):715–727.

Stern, D. N. (1977). *The First Relationship: Infant and Mother.* Harvard Univ Pr.

Stern, D. N. (1985). *The Interpersonal World of the Infant: A View from Psychoanalysis and Developmental Psychology.* Basic Books.

Stern, D. N. (2010). *Forms of Vitality Exploring Dynamic Experience in Psychology, Arts, Psychotherapy, and Development.* Oxford Press.

Suk, H. J., Jeong, S. H., Hang, T. H., and Kwon, D. S. (2009). Tactile sensation as emotion elicitor. *Kansei Eng. Int.*, 8(2):147–52.

Takayama, L., Dooley, D., and Ju, W. (2011). Expressing thought: improving robot readability with animation principles. *6th ACM/IEEE Int. Conf. Human-Robot Interact.*

Tamietto, M., Cauda, F., Celeghin, A., Diano, M., Costa, T., Cossa, F. M., Sacco, K., Duca, S., Geminiani, G. C., and de Gelder, B. (2015). Once you feel it, you see it: insula and sensory-motor contribution to visual awareness for fearful bodies in parietal neglect. *Cortex.*, 62:56–72.

Tomkins, S. S. (1993). *Affects, imagery, consciousness: The positive affects.* Springer. New York.

Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., and Rizzolatti, G. (2001). I know what you are doing. a neurophysiological study. *Neuron.*, 31(1):155–65.

Vannucci, F., Di Cesare, G., Rea, F., Sandini, G., and Sciutti, A. (2018). A robot with style: Can robotic attitudes influence human actions? *IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids).*

Varcin, K. J., Grainger, S. A., Richmond, J. L., Bailey, P. E., and Henry, J. D. (2019). A role for affectivity in rapid facial mimicry: An electromyographic study. *Soc Neurosci.*, 14(5):608–617.

Vogt, B. A. and Pandya, D. N. (1987). Cingulate cortex of the rhesus monkey: Ii. cortical afferents. *J Comp Neurol.*, 262(2):271–289.

Wallach, W. and Allen, C. (2008). Moral machines: teaching robots right from wrong. *Oxford University Press.*

Willis, J. and Todorov, A. (2006). First impressions: making up your mind after a 100-ms exposure to a face. *Psychol Sci.*, 17(7):592–8.

Wong, A. L., Haith, A. M., and Krakauer, J. W. (2015). Motor planning. *Neuroscientist.*, 21(4):385–98.

Zeidman, P., Jafarian, A., Corbin, N., Seghier, M. L., Razi, A., Price, C. J., and Friston, K. J. (2019a). A guide to group effective connectivity analysis, part 1: First level analysis with dcm for fmri. *Neuroimage.*, 200:174–190.

Zeidman, P., Jafarian, A., Seghier, M. L., Litvak, V., Cagnan, H., Price, C. J., and Friston, K. J. (2019b). A guide to group effective connectivity analysis, part 2: Second level analysis with peb. *Neuroimage.*, 200:12–25.