



**SAPIENZA**  
UNIVERSITÀ DI ROMA

# Integrating Sensory-Motor and Social Dynamics: Multisensory Processing and Social Factors in Interpersonal Interactions

**Final Dissertation**  
**International Doctor of Philosophy**  
**Department of Psychology**  
**Cognitive, Social and Affective Neuroscience**

**Giovanna Cuomo**  
**XXXVII Cycle**

**Supervisor**  
Matteo Candidi

**External Reviewers**  
Francesca Ferroni  
Ivan Patané

A.A. 2021-2024

*A chi ci ha creduto fin dall'inizio e fino alla fine.*

*Ai miei genitori, che lo hanno fatto anche senza capirmi sempre, perché erano fieri di me.*

*Alle mie sorelle, con cui sono cresciuta, che non sanno di avermi dato la forza.*

*Alla mia persona, per esserci stato sempre, soprattutto quando non era facile.*

*Alle mie guide nella ricerca, che rappresentano l'ideale a cui aspiro.*

*Alle mie amicizie e ai miei colleghi, che mi hanno sostenuta e accompagnata giorno dopo giorno.*

**SUPERVISOR**

Prof. Matteo Candidi

Social and Cognitive Neuroscience Laboratory, Department of Psychology, Sapienza University of

Rome, Italy

## **EXTERNAL REVIEWERS**

PhD. Francesca Ferroni

Social Cognitive Neuroscience Lab, Department of Medicine and Surgery - Unit of Neuroscience,  
University of Parma, Italy

The doctoral thesis by Giovanna Cuomo provides an in-depth investigation into the intricate relationship between interpersonal motor interactions and multisensory processing. By focusing on the integration of visual and tactile stimuli, motor control during joint actions, and the influence of conversational contexts on interpersonal representation processes, the research offers valuable insights into the mechanisms underlying human interactions. The work is characterized by good originality, a robust methodological approach, and significant theoretical and practical contributions.

PhD. Ivan Patané

DipSCo, Department of Psychology and Cognitive Sciences -

University of Trento, Italy

It's been a pleasure to evaluate this doctoral thesis, and I would like to thank the candidate and her supervisor for entrusting me with this task. The work demonstrates a solid level of academic rigor and intellectual depth, offering valuable insights into the social neuroscientific field.

The thesis is well-structured, with both the introductory and experimental sections contributing meaningfully to the overall work. The candidate has successfully navigated a wide array of theoretical frameworks, ranging from multisensory integration to racial bias, presenting them with clarity and precision. The ability to deliver the most relevant literature while avoiding unnecessary detail ensures that the narrative remains focused and coherent. This allows readers to get the necessary theoretical foundation without being overwhelmed, effectively setting the stage for the subsequent experimental work.

The experimental component is noteworthy for its methodological diversity and the robustness of the results. The candidate demonstrates fine skill in designing and implementing a wide range of experimental approaches, showcasing both technical competence and versatility. The rigor with which the experiments were conducted highlights the candidate's mastery of diverse methodologies and her capacity to engage with complex scientific questions at a high level.

While the thesis is of high quality, I provide a few minor suggestions and discussion points below. These are intended to encourage further reflection and may serve as topics for elaboration during the defence. They are not critical but rather opportunities to expand on certain aspects .

## **Abstract**

Human interaction with others requires the integration of sensory-motor and cognitive processes. In social contexts, coordinating actions with a partner depends on mechanisms such as prediction and adaptation, which are influenced by multisensory and motor information. To create unified perceptual experiences and interact efficiently, the brain must integrate inputs from various sensory modalities. Conversely, social perception and contextual factors also influence our sensorimotor experience.

This thesis examines whether interpersonal motor interactions affect multisensory processing, with a specific focus on the integration of visual and tactile stimuli, and motor control during or after joint actions. It also explores how social variables, such as conversational contexts, shape interpersonal representation processes after linguistic interaction. The findings discussed in the present work highlight how dynamic social, linguistic and sensorimotor information modulate perception and action, offering insights into the mechanisms that underpin human interactions.

## List of Publications Associated with the PhD

Fini, C., Era, V., **Cuomo, G.**, Falcinelli, I., Gervasi M.A., Candidi, M., Mazzuca, C., Liuzza M.T., Winter B., Borghi, A.M. (preprint). Digital Connection, Real Bonding: Brief OnlineChats Boost Interpersonal Closeness Regardless of the Conversational Topic. Heliyon

Fini, C., Falcinelli, I., **Cuomo, G.**, Era, V., Candidi, M., Tummolini, L., Mazzuca, C. & Borghi, A. M. (2023). Breaking the ice in a conversation: abstract words prompt dialogs more easily than concrete ones. *Language and Cognition*, 15(4), 629-650.

## List of Publications Not Associated with the PhD Supervisor

**Cuomo, G.**, Maglianella, V., Ghanbari Ghooshchy, S., Zocolotti, P., Martelli, M., Paolucci, S., Morone, G. & Iosa, M. (2022). Motor imagery and gait control in Parkinson's disease: Techniques and new perspectives in neurorehabilitation. *Expert Review of Neurotherapeutics*, 22(1), 43-51.

# Table of Contents

1		
General Introduction		10
1.1. Perception and Action in Interpersonal Contexts		11
1.2. Body and Space Representation During Interpersonal Interactions		15
1.3. Multisensory Integration and Peripersonal Space Representation		18
1.4. Modulations of Cross-Modal Effect and the PPS		21
1.4.1. Action Planning and Execution		21
1.4.2. Social Factors		22
1.5. Social Factors Influence on Interpersonal Interactions		25
1.5.1. From Sensorimotor Contingencies to Social Cognition		26
1.5.2. From Communicative Gestures to Language		28
1.6. Thesis Contributions and Hypotheses		30
2		
Unraveling Multisensory Integration in Interactive Contexts		33
2.1. On-line Interpersonal Motor Interactions Modulate Visuo-Tactile Integration Mechanisms		34
2.1.1. Introduction		34
2.1.2. Materials and Methods		36
2.1.3. General Conclusion		72
2.2. Dyadic interactions based on multiple sensorimotor modalities modulate the behavioral and electrocortical markers of interpersonal cross-modal processing		76
2.2.1. Introduction		76
2.2.2. Materials and Methods		77
2.2.3. Results and Discussion		83
2.3. Toward a “Social Rubber Hand”		90
2.3.1. Introduction		90
2.3.2. Material and Methods		92
2.3.3. Data Analyses and Results		101
2.3.4. Discussion		108
2.4. The Hidden Cost of Outgroup Coordination: Action Delays in Joint Actions.		112
2.4.1. Introduction		112
2.4.2. Materials and Methods		114
2.4.3. Results		124
2.4.4. Discussion		140
2.5. Interim Summary		146

3		
Communication Dynamics in Interpersonal Contexts		148
3.1. Breaking the ice in a conversation: abstract words prompt dialogs more easily than concrete ones.		149
3.1.1. Introduction		149
Study 1		152
Method		152
Data analysis		155
Study 2		168
Method		168
Data analysis		169
3.1.2. Discussion		174
3.2. Digital Connection, Real Bonding: Brief OnlineChats Boost Interpersonal Closeness Regardless of the Conversational Topic		178
3.2.1. Introduction		178
Experiment 1		182
Experiment 2		195
3.2.2. General Discussion		202
3.4. Summary		208
4		
General Discussions		209
Appendix		213
A1. Overview of the Publication Status of Chapters in this Thesis		213
A2. Preregistration “ Processing Abstract and Concrete Concepts: the Case of Aphantasia”		214
A2.1. Introduction		214
A2.2. Matherials & Methods		215
A2.3. Analysis plan		222
References		224

# 1

## General Introduction

Everyday life requires executing movements, actions, and interactions with objects and other individuals. To effectively engage with the environment, sensory-motor and cognitive systems are involved. In social contexts, the ability to couple executed actions with those of a partner is essential for achieving shared goals during coordination. Mechanisms such as prediction and adaptation to others' movements and sensorimotor information play a key role in shaping social behaviour alongside various social factors that may influence these interactions. Moreover, during typical interpersonal interactions, information from multiple sensory modalities is often simultaneously available. For efficient responses, multisensory integration processes are essential to create a unified perceptual experience.

The present thesis investigates how individuals predict, process, and adapt to others' behavior during interpersonal motor interactions, focusing on the role of (low-level) multisensory processing and integration of (higher-order) social variables in shaping sensorimotor processes. Specifically, it explores how different types of interactions with a partner affect the processing of multisensory, visuo-tactile, stimuli and motor control (Chapter 2), as well as interpersonal representations (Chapter 3). The main objectives of the work described in the thesis are to: 1) Examine how multisensory representations, particularly visual and tactile stimuli, are integrated and influenced during joint action tasks (see 2.1), and how these processes are modulated after an interaction at both behavioural and neurophysiological (i.e., electroencephalography, EEG) levels (see 2.2); 2) Investigate how the temporal synchronicity between visuo-tactile information during motor interactions influences multisensory processing and perceptual outcomes (see 2.3); as well as the impact of individual differences such ethnicity (see 2.4), on sensorimotor and cognitive processes during interpersonal interactions. 3) Explore the impact of linguistic interactions, including abstract and concrete conversational contexts, on interpersonal representations (see 3.1 and 3.2.).

The first chapter provides an overview of the latest advancements in research on sensorimotor body and space perception, interpersonal sensorimotor interactions, multisensory integration processes and social dynamics. Following this introduction to the state of the art, the subsequent chapters present behavioural and neurophysiological findings that contribute to a deeper

understanding of perceptual and cognitive mechanisms in social interactions. These findings elucidate the role of interpersonal interactions in shaping multisensory integration, motor control and social cognition, offering critical insights that will underpin future theoretical developments and practical applications.

## **1.1. Perception and Action in Interpersonal Contexts**

In our everyday lives, we are continually engaged in performing or observing actions that are closely intertwined with our ability to perceive and interact with the world around us. This interaction between motor processes and perception has deep roots in psychological and neuroscientific theories. According to Sperry (1952), sensory and motor processes are integrated in a continuous feedback loop that shapes our understanding and interactions with the environment, laying the groundwork for later theories on motor-perception integration. This theoretical framework suggests that action-based processes do not merely follow perception but actively modulate it, influencing how sensory information is processed. Building on this, later studies showed that motor intention could actively shape perception. For instance, Craighero et al. (1999) demonstrated that preparing a grasping movement can influence the detection and discrimination of visual stimuli. In their study, participants were asked to grasp a bar after being presented with a visual stimulus that either matched or did not match the orientation of the bar. The results indicated that movement preparation improved the detection of visual stimuli presented congruently with the object's properties, showing how motor processes influence perception. Similarly, De' Speriati and Stucchi (1997) found that right- and left-handed participants used internal motor simulations of their dominant hand's movement to differentiate between observed screwing and unscrewing actions. In related studies, Gentilucci et al. (1998a, b) showed that people more easily recognized common hand postures, typically used in object interaction, than unusual postures, even when the latter provided more visual information for handedness recognition. These findings highlight how motor knowledge supports perceptual tasks.

The theory of 'common coding' (Prinz, 1990) provided another conceptual model to explain the connection between perception and motor processes, suggesting that perceptual and motor representations share overlapping codes. Substantial evidence supports a common neural framework between action and perception in humans (Viviani, 2002; Fadiga et al., 2000), indicating the possibility in social contexts for a direct translation between observing and performing an

action. For example, mirror neurons in the monkey's ventral premotor area (F5) fire both when the monkey performs a goal-directed action and when it observes the same action being executed by another individual (Rizzolatti et al., 1996). Over the years, mechanisms such as mirror neuron system have given insight into the abilities to assume a first-person perspective on the actions (Rizzolatti & Craighero, 2004), sensations (Whicker et al. 2003; Gallese et al., 2004), emotions (Hatfield et al., 1993) and intentions of others (Rizzolatti & Sinigaglia, 2010) and form interpretations of others' mental states from a third-person perspective (Firth & Firth, 2006; 2010). This framework emphasizes that in social contexts, when interacting, not only individuals are able to mentalize the other's mind, but lower functional motor mechanisms, as the mirror system, seem to be at the base of how we experience others (cfr. Gallese et al., 2004). This indicates that to truly grasp the behavioural and neural mechanisms that underpin cognition, it is important to shift the focus to how perception and action operate in social environments in what has been referred to as 'second-person neuroscience' (Schilbach et al., 2013).

Action is related to perception also through selective attention, as stated in the 'premotor theory of attention' (Rizzolatti et al., 1994). This theory posits that a single control system governs both attention and action, especially in spatial contexts, and that is essential for pragmatic reasons (e.g., without spatial awareness, it would be impossible to efficiently act upon or interact with the environment). Looking at this perspective in a social framework, the premotor theory of attention finds support in the case of social learning and joint attention mechanisms during interpersonal exchanges. The ability to selectively align one's attention with another person's focus on objects and events of mutual interest provides a basic mechanism of a common cognitive framework from early development (Tollefsen, 2005; Frischen & Tipper, 2004). Humans utilize a variety of social signals, such as gaze, gestures, and posture, to indicate which object or event is being attended to. Indeed, mirroring abilities are also involved in these attentional processes: motor imitation plays a crucial role in learning and communicating the meaning of objects, events, and movements. The active participation in joint attentional processes facilitates the creation of a shared focus in interaction to form a perceptual common space (Sebanz et al., 2006), a shared mental representation, aligning two individuals' perspectives on the same experience, supporting the initiation of coordinated actions, and facilitating ongoing coordination during tasks.

However, sudden environmental changes might cause the failure to adjust to the other focus or action, unless we are able to predict them. This could be related to the concept of 'forward models', the brain mechanism that predicts the consequences of motor actions, connecting motor control to cognitive processes, two processes that might underlie our abilities to imitate and

coordinate with others (Wolpert et al., 1995; Miall, 2003). Indeed, imitative actions benefit from the overlap of motor representations between partners, allowing anticipatory alignment of actions based on internal simulative mechanisms for one's motor control while simultaneously simulating the others' actions—potentially using the same internal models for both (Blakemore & Decety, 2001; Wolpert et al., 2003; Pezzulo et al., 2013; Dindo et al., 2011; Pezzulo, 2011). In this framework, imitation builds upon shared attention by translating the observed actions and intentions of others into one's own motor representations, enabling a deeper understanding and synchronization of actions with a conspecific (Meltzoff & Decety, 2003; Iacoboni et al., 2005). Individuals seem to treat the co-actors actions similarly to their own during action planning, potentially forming shared task representations when interacting to achieve a common goal (Sebanz & Frith, 2004).

Furthermore, joint attention and imitation seem to rely on the ability to infer another person's action goals (Bekkering et al., 2000; Wohlschläger et al., 2003). Neurons located on the lateral surface of the inferior parietal lobule, demonstrate selectivity not only for the movement currently being executed, but also for the subsequent movements in a sequence, indicating their involvement in representing the overall objective of a chain of movements (Fogassi et al., 2005; Tunik et al., 2007). Indeed, successful interactions seem to require understanding others' actions and intentions, to anticipate their next moves, and predict the outcomes (Verfaillie & Daems, 2002). Planning one's actions in interpersonal interactions involves accounting for the timing, type and goal of the others' movements, foreseeing and adjusting to their actions. The capability to merge one's movements with those of another and to align individual sub-goals within a unified motor strategy can be conceptualised as a motor shared goal representation. This process is guided by a joint goal representation that predicts the combined outcome of both participants' actions (Sebanz et al., 2006). Evidence suggests that motor resonance in imitation facilitates these predictive abilities (Wilson & Knoblich, 2005; Kilner et al., 2004; Miall, 2003). While imitation is beneficial in many social situations, such as observational learning, it can be counterproductive in everyday dyadic interactions requiring complementary actions to reach shared objectives successfully.

Understanding what someone else focuses on in a given context provides important clues about their intentions and can trigger imitative, but also complementary actions in the observer (Richardson & Dale, 2005). Achieving complementary movements is possible when internal motor representations triggered by observation (Iacoboni et al., 1999) are inhibited by a shared goal representation (Fogassi et al., 2005), that predicts the combined outcome of both participants' actions (Sebanz et al., 2006). This inhibition allows to choose suitable actions to accomplish joint objectives, allowing an individual to carry out actions that differ from those observed (Meltzoff &

Decety, 2003; Chaminade et al., 2005). In contrast to imitation, complementary actions demand a mismatch in either time or space between participants' movements and objectives (Hasson & Frith, 2016). Neuroimaging findings indicate that both imitation and complementary interactions might engage a distributed fronto-parietal network (Newman-Norlund et al., 2007). For instance, the anterior intraparietal sulcus (aIPS) appears to play a role in coordinating shared goals by predicting and integrating complementary actions between individuals (Sacheli et al., 2015a; Era et al., 2018). Meanwhile, coordination abilities necessary for synchronous actions seem to be related to the activity of the primary motor cortex (Novembre et al., 2014) and dorsal premotor cortex (Hadley et al., 2015). Complementary actions, however, might call for greater neural resources to integrate and adapt one's actions with others (Kokal et al., 2009). Indeed, successful interactions likely relies not only on anticipating the partner's behaviour but also on understanding what the task requires (Sebanz et al., 2005; Konvalinka et al., 2010; Gallotti et al., 2017). In this case, action simulation could involve additional steps for transforming the perspective from the observer to the others' frame of reference -often referred to as perceptual (Johnson & Demiris, 2005) or visuospatial perspective taking (Zacks & Michelon, 2005) – leading evidence for the connection between space attention and action processes.

In summary, by examining different types of evidence that relate perception to action, it is clear that different neural mechanisms could support social interaction, with certain brain regions and functions potentially contributing depending on the objective and type of interpersonal interactions. It was possible to distinguish between two critical aspects of motor behaviour: the anticipation of others' movements (Knoblich & Jordan, 2003) and the mental representation of shared goals (Butterfill, 2012). Perceptual, motor, and cognitive functions allow individuals to successfully act together. Shared intentionality, sensorimotor coupling, and synchronized movements between co-actors, are suggested to enable them to influence one another (Sebanz et al., 2006). As stated, the use of predictive mechanisms potentially allow individuals to anticipate both their own actions and those of their partner to integrate this information into a shared action plan (Sebanz & Knoblich, 2009). This conceptual framework suggest that while interacting to pursue a common goal, the action could be divided into distinct sub-tasks, with each individual responsible for executing their part to ensure the completion of the collective activity. This action coordination between two or more individuals as a cohesive unit in space and time to achieve a shared goal is referred to as 'joint action' (Sebanz et al., 2006; Konvalinka & Roepstorff, 2012). In this process, participants can be conceptualized as coupled dynamical systems that mutually predict and adapt to one another's actions (Kelso et al., 2013) at behavioural (Richardson et al., 2007), physiological (Muller &

Lindenberger, 2011; Mitkidis et al., 2015) and neural levels (Konvalinka & Roepstorff, 2012; Novembre et al., 2017). To capture this dynamic, it could be useful to introduce the concept of a 'Shared Action Space' (Pezzullo et al., 2013), better discussed in next sections, suggesting that individuals engaged in joint action might incorporate their partner's movements and intentions into their own spatial representation, allowing for the coordination of both agents' motor plans.

## **1.2. Body and Space Representation During Interpersonal Interactions**

We perceive and move in space through multiple function-specific representations held within different cortical areas which in turn may be responsible for the control of movements of different body parts (Rizzolatti et al., 1994; Gross & Graziano, 1995; Colby & Duhamel, 1996). Acting within an environment not only involves the coding and interpretation of our surroundings, but also entails a robust representation of ourselves. Inputs in different sensory modalities form a representation of one's own body known as the 'body schema'. Indeed, the term body schema refers a constantly updated multimodal neural representation of the body that adapts during movement and motor learning (Botvinick, 2004; Head & Holmes, 1911). This concept is distinct from another representation of the body, known as the 'body image', which is referred to as a conscious representation or beliefs regarding one's body (Gallagher & Meltzoff, 1996) and it is used to indicate an implicit knowledge of our body's form, configuration, and how it influences touch, vision and movement (Graziano & Botvinick, 2002).

As individuals grow or acquire new motor skills, the body schema adjusts to reflect changes in action capabilities. This schema is thought to extend to incorporate objects relevant to the action, such as tools. In support to this hypothesis, studies investigating the receptive fields of neurons in the intraparietal area of monkeys, which respond to both somatosensory and visual stimuli, are found to expand and include the full length of a tool after its use (Maravita & Iriki, 2004; Iriki et al., 1996). In other words, tool use appear to modify the body schema by extending the body's internal representation, as proved also in human (Arbib et al., 2009; Cardinali et al., 2009; Maravita et al., 2002). For instance, patients with hemispatial neglect show deficits in far space when using tools, indicating that tool use affects spatial representation beyond immediate reach (Berti & Frassinetti, 2000). These evidences indicate that individuals perceptual system adapts the body's representations to optimize the sensorimotor transformations necessary for effective actions, such as reaching and

grasping objects and predicting their sensory consequences (Wolpert et al., 1995; Pouget et al., 2002; Miall, 2003). This close relationship between body, action and spatial representation suggests that the brain encodes body and external spatial locations ensuring flexibility and accuracy in movement (Gross & Graziano, 1995). The space encoding of the body and its' surroundings is based on the integration of different sensory and proprioceptive inputs (Maravita et al., 2003a). These processes are suggested to involve a network of brain regions related to the dorsal stream, especially the posterior parietal cortex (Colby & Goldberg, 1999; Ferraina et al., 2009), the premotor cortex (Graziano et al., 1994; Rizzolatti & Luppino, 2001), and the cerebellum (Wolpert et al., 1995).

Specifically, based on evidence from brain-damaged patients, Goodale and Milner (1992) introduced a distinction between two modes of perception: one aimed at object recognition (perception-for-identification) and the other aimed at guiding action (perception-for-action). The authors later supported this division by linking object identification to the ventral (occipito-temporal) stream and action guidance to the dorsal (occipito-parietal) stream (Milner & Goodale, 2008). Patients with damage in the superior parts of the posterior parietal cortex are reported to frequently struggle to use information about an object's size, shape, orientation, and location to guide the hand posture, finger positioning, or limb movement when attempting to grasp it (Goodale & Milner, 1995; Goodale et al., 1994). However, these individuals can often still recognize and describe the very objects that they cannot effectively grasp (Milner & Goodale, 1995; Goodale & Milner, 1992). Conversely, some patients display the opposite pattern, able to perform grasping actions but impaired in object recognition (Goodale & Meenan et al., 1994; Goodale & Milner, 1992). In healthy individuals, this distinction has also been proposed as a potential explanation for the perceptual and action-related differences in object size judgments (Aglioti et al., 1995). This pattern of dissociation is supposed to reflect the distinct computations performed by these two visual pathways in the brain (Milner & Goodale, 1995). The ventral pathway relies on relative metrics and an object-centered frame of reference, while the dorsal pathway encodes spatial information egocentrically, translating distances in a way that is directly relevant for action (Ganel et al., 2008). The dorsal stream is responsible for organizing the space around us for perceptual purposes and enabling effective interaction with the environment through action (Goodale & Milner, 1992; Rizzolatti et al., 1997) by converting visual information into a more relevant body-centered coordinate frame (Jeannerod & Biguer, 1989).

It has been proposed that when we move with a partner as well, our body schema would adjust as an effect of 'social entanglement' to include and incorporate the other's body movement

(Soliman et al., 2015). This phenomenon has been referred to as ‘joint body schema’ (JBS). Soliman and colleagues (2015) proposed that joint actions—beyond cognitive mechanisms like joint attention, action observation, and intention understanding—may also depend on overlapping body representations of the self and the other. This JBS, i.e., the integration of visual stimuli seen on the hand of our partner with tactile stimuli felt on one’s own hand, seems to persist beyond the task itself once established during an interpersonal interaction (Soliman et al., 2015). Such shared schema is thought to emerge when individuals must coordinate their actions to achieve shared goals, as in joint action tasks that might lead to the integration of body representations between co-actors to coordinate their movements better. By gaining a better understanding of their partner's potential movements, individuals can adjust their own actions to better complement their partner's motions. Taking a step back to the discussed distinct computations acted by the ventral and dorsal visual streams, this JBS transformation when individuals are spatially aligned could reflect an egocentric-to-egocentric frame (Becchio et al., 2013). Meaning that when two individuals are spatially aligned, they share a similar viewpoint or are oriented in the same direction during an interaction, they could create a joint body schema using egocentric (self-centered) frames to interact efficiently. However, when viewpoints are misaligned, integrating two different frames of reference could be harder, leading to an object-centred remapping to coordinate actions effectively (Bockler et al., 2011). In the latter case, the individual action spaces are merged into what can be referred to as a joint action space, where the agents' perspectives and orientations remain unaligned, but they coordinate actions by relying on a shared spatial framework rather than a common viewpoint (Pezzullo et al., 2013).

This shared space is referred to as Shared Action Space (SAS). Building on the notion that interpersonal interactions can reshape the space surrounding the body by expanding to the space of a partner, individuals integrate the operational spaces of their partners into their spatial representations encompassing areas where direct interaction occurs, even if beyond their immediate reach, allowing for more efficient complementary or simultaneous joint actions (Pezzullo et al., 2013). Individuals naturally rely on their own body schema to respond to sensory information near or around the bodies of others (Brozzoli et al., 2013; Decety & Sommerville, 2003; Thomas et al., 2006). Building on this notion, physiological evidence supports the existence of integrated spatial representations, sensitive to the spatial contexts in which another's actions occur (Caggiano et al., 2009), reminding to mirror-like mechanisms. Furthermore, mirror neurons are part of a broader *action observation network* (AON), which includes regions in the parietal, premotor, and occipitotemporal cortices processing various types of information regarding others’ movements and

goals, representing the possible sensory and motor foundations supporting effective joint actions (Kessler et al., 2006; Biermann-Ruben et al., 2008; Grafton, 2009; Neal & Kilner, 2010). In a shared interpersonal context, multisensory and motor integration could be considered particularly relevant for their intrinsic avoidance and approaching function between the body and the environment. Particularly, visuo-tactile neurons are significant, due to the essential role of vision in directing our actions and predicting their tactile outcomes in daily life. Through visuo-tactile integration (VTI) effects these bimodal neurons are proved to respond to visual stimuli not just near the individual, but also related to another person's body (Brozzoli et al., 2013). These spatial sensory-motor neural mechanisms indicate an altered perception of the space when it is shared.

### **1.3. Multisensory Integration and Peripersonal Space Representation**

Different studies have identified both in monkeys (Duhamel et al., 1998; Hyvarinen & Poranen, 1974; Rizzolatti et al., 1981a, 1981b) and humans (Serenio & Huang, 2006) different body-parts centered bimodal neurons (e.g., coding for stimuli in two sensory modalities) that seem to be the neurophysiological basis of the integration between stimuli in different sensory modalities. Focusing on visuo-tactile mechanisms, which encompass sensory and motor-related functions, they serve as a multisensory-motor interfaces to support effective interactions with objects and the environment. Indeed, certain visuo-tactile neurons in monkeys exhibit a dynamic nature in their visual receptive fields. There are notable differences between the receptive fields (RFs) of tactile and visual neurons: tactile RFs are fixed to specific body parts, while visual RFs surrounding the same body parts can respond to objects approaching the body part within approximately 30 cm (Rizzolatti et al., 1981a, b; Gentilucci et al., 1988; Graziano & Gross, 1996; Fogassi et al., 1996). These neurons, termed 'peripersonal neurons' (Rizzolatti et al., 1981a, b; Gentilucci et al., 1988), were identified based on their sensitivity to stimuli appearing close to the body's surface, specifically within the peripersonal space (PPS). Fogassi et al. (1996) found that the visual receptive fields of F4 visuo-tactile neurons can expand in depth when a visual stimulus approaches the cutaneous receptive field at varying velocities, specifically between 20 and 80 cm/s. This characteristic is crucial for both defensive mechanisms and the preparation and execution of actions directed toward objects within proximity, again within the PPS. Additionally, research by Iriki et al. (1996) revealed that after monkeys were trained to utilize a rake as a tool to access food pellets located outside their normal reach, certain neurons in the post-central gyrus (with some extension

into the intraparietal sulcus) began to demonstrate visual activity. While some authors have raised questions about these last findings (Holmes & Spence, 2004), it was observed that these visual responses were more pronounced during active tool usage compared to passive interactions, and visual receptive fields appeared to stretch toward the rake's tip. Remarkably, a few minutes following active tool use, these visual receptive fields seemingly returned to their original size. This indicates that the dynamic properties of visual receptive fields may be closely linked to the performance of specific motor actions, that can impact and re-shape PPS representations (Rizzolatti et al., 1998).

These findings suggest that the dynamic adjustment of multisensory receptive fields and perceptual representations is directly linked to goal-directed actions, not just responding to stimuli close to the individual's body parts, aligning with the concept of a shared coding for perception and action (Prinz, 1997, 1990). This is in line with the notion that such bimodal neurons are mainly located within areas involved in sensorimotor integration in monkeys, such as the inferior area 6 (Gentilucci et al., 1988; Rizzolatti & Gentilucci, 1988; Rizzolatti et al., 1987, 1988, 1997), parietal areas 7b (Hyvärinen, 1981; Hyvärinen & Poranen, 1974; Hyvärinen & Shelepin, 1979; Leinonen, 1980; Leinonen & Nyman, 1979), and the putamen (Crutcher & DeLong, 1984). This neuronal organization support the close relationship between space coding and sensorimotor processing: objects located within the reaching distance increase their salience as a function of the possible interactions that we can establish, supporting a function in effective guidance of the motor response (Graziano & Gross, 1993; Graziano, 1999; Rizzolatti et al., 1997, 1998; Maravita et al., 2003a, b; Maravita, 2006). Interestingly, visuo-tactile neurons within area 7b and the ventral section of the intraparietal sulcus display mirror-like properties when observing, independently from the observing position, another's body-part approached by a visual cue (Ishida et al. 2010). This finding might suggest a role for these regions not only to interact with nearby objects, but also to understand and code for conspecific interactions with the environment.

In humans, several studies support these findings for visuo-tactile stimuli presented near-body in both neuropsychological conditions (Làdavvas, 2002; Làdavvas & Farnè, 2004; Legrand et al., 2007) and healthy population (Brozzoli et al., 2010, 2009; Spence et al., 2004, 2008). Patients affected with 'extinction' show impaired detection of stimuli presented contralaterally to their brain lesion (in the right parietal lobe) if they are presented synchronously with a stimulus in the same sensory modality on the same side as the lesion (Bender & Feldman, 1951; Bender, 1952; Brozzoli et al., 2006). This phenomenon is thought to occur due to competition for awareness between the impaired and intact spatial representations (Jacobs et al., 2011). The extinction can also be crossmodal, it has

been found that visual cues near the hand contralateral to the lesion are extinguished by a tactile stimulus on the other hand (Mattingley et al., 1997) and that the extent of crossmodal extinction experienced by the patients' varies as a function of the stimuli distance from their body (di Pellegrino et al. 1997). This is also true for face-centred stimulations (Farnè et al. 2005a). Furthermore, these patients showed crossmodal extinction for extrapersonal space after a tool use of 5-10 minutes to retrieve or act upon distant objects (Farnè & Làdavas, 2000; Berti & Frassinetti, 2000; Bonifazi et al., 2007; Farnè et al., 2005b, 2007; Maravita & Iriki, 2004).

These findings are in line with studies on healthy subjects, investigated through crossmodal congruency task (Pavani & Castiello, 2004; Spence et al., 2004, 2008), where participants respond more quickly and accurately when both the tactile and visual stimuli are aligned at the same elevation with respect to their fingers, a phenomenon known as the crossmodal congruency effect (CCE; Spence et al., 2000). Notably, the intensity of the CCE increases when the visual stimuli are nearer to the hand (e.g., within the PPS) compared to when they are presented farther away (Pavani et al., 2000; Maravita et al., 2003a; Macaluso & Maravita, 2010) and extend to far space after tool use (Maravita et al., 2002; Holmes et al. 2004, 2007a, b). The locations for humans' bimodal neurons coding for these VTI effects correspond to the ones found in monkeys, specifically they are reported to be in the anterior section of the intraparietal sulcus and the ventral premotor cortex (Bremmer et al., 2001; Makin et al., 2007; Serino et al., 2007). These areas have been further confirmed in studies investigating multisensory information around the face (Bremmer et al., 2001) and hand regions (Serenò & Huang, 2006), although different body parts, such as the face or hands, have spatially distinct but overlapping visuo-tactile topographical representations (Huang et al., 2012). Recent research using intracranial electroencephalography demonstrated that also certain brain regions, such as the post-central gyrus and insula, integrate multisensory inputs with proximity modulation (Bernasconi et al., 2018) as in non-human primates (Avillac et al., 2007). These results imply that, like in monkeys, the frontal and parietal regions associated with multisensory processing in humans are part of a larger fronto-parietal network responsible for sensory-motor processes that enable interactions with the environment and dynamically modulate individuals' space representation mechanisms (Rizzolatti et al., 1997, 1998).

## **1.4. Modulations of Cross-Modal Effect and the PPS**

### **1.4.1. Action Planning and Execution**

The VTI mechanisms influence and are influenced by the representations of the space surrounding our bodies, literally the PPS (di Pellegrino & Ladavas, 2015; Patané et al., 2017; Cléry & Hamed, 2018; Bufacchi & Iannetti, 2018; Serino, 2019). This representation, which relies on the previously cited multisensory neurons (Rizzolatti et al., 1981a, b; Gentilucci et al., 1988; Graziano & Gross, 1996; Fogassi et al., 1996), serves as a sensorimotor interface that enables interaction with the environment, particularly during action planning and execution such as, for example, grasping and reaching movements. The perception of whether external objects are located inside or outside the PPS is influenced by the characteristics of the body's movements and the location of its various parts. These multisensory mechanisms underlying PPS representations are believed to improve voluntary actions like reaching for and grasping nearby objects (Brozzoli et al., 2009).

However, the integration of different sensory modalities in PPS can be modulated by individual's action intentions or specific interacting goal (Brozzoli et al., 2010; Patané et al., 2019; Senna et al., 2019). For instance, recent behavioral research indicates that multisensory integration effects happening in the near-body space are dynamically extended to portions of space far from the body in anticipation of voluntary body movements toward these regions (Brozzoli et al., 2009, 2010, 2012; Makin et al., 2012). Indeed, when participants are asked to discriminate a tactile stimulus delivered to a finger of the hand they will move to reach-and-grasp an object while a visual cue is presented in either a congruent or incongruent position relative to the final movement location on an object, VTI effects become evident even in the case of visual cue outside of their PPS (Brozzoli et al., 2010). Moreover, similar effects have been observed for peri-trunk space (Canzonieri et al., 2012) and whole-body movements (Noel et al., 2014).

Patané and colleagues (2019) demonstrated that cross-modal visuo-tactile integration can be modulated by action from its planning phase under specific conditions, as when the orientation of the object to be grasped was randomized for each trial, introducing uncertainty. Participants received the tactile stimulation during different phases of the movements, from its planning to the grip on the object. However, they could only fully plan their action (in terms of object orientation) after seeing the object, that constituted the go signal of the movement. The interaction caused by the orientation of the visual cues relative to the action on tactile discrimination suggested that, even in the planning phase, there was an early reprocessing of the relationship between visual and tactile

stimuli, suggesting a remapping of the PPS representation. The timing in which these VTI effects were found during the action execution and their linear strengthening further suggests that continuous monitoring of the action and the action space occurs throughout the entire movement. This appeared also evident at the kinematic level (Patanè et al., 2019).

The modulation of VTI effects during action planning and execution was further examined in situations where sequences of actions were required during action execution. Senna and colleagues (2019) explored how combinations of actions and their plausibility affected the intensity of the CCE. Participants were tasked not only with grasping actions over a cylindrical candy-like object that displayed task-irrelevant visual cues, but also with bringing it toward themselves. The plausibility of the action was determined by the orientation of the object (either horizontal or vertical), which varied on a trial-by-trial basis, similar to the setup used by Patanè et al. (2019). Reaction times showed that the strongest CCE occurred only when participants were holding the cylinder with a plausible action—that is, when the candy was positioned horizontally, allowing it to be easily brought to the mouth without additional reorientation, unlike the vertical position. These findings differ from earlier studies (Brozzoli et al., 2009, 2010; Patanè et al., 2019) that reported a continuous, linear updating of the relationship between visual distractors and tactile stimuli, beginning at the action planning phase and extending throughout movement execution and approach to the target object (Patanè et al., 2019). The findings from Senna et al. (2019) suggest that the remapping of the PPS area may occur at different stages of the movement chain execution, depending on the specific action aims and the plausibility of those actions. These results enhance our understanding of how peripersonal space functions to support and guide voluntary actions aimed at specific goals.

#### **1.4.2. Social Factors**

In light of the describe research on VTI effect in the space of another, it is reasonable to believe that PPS representation plays a crucial role in physical interactions not only with object in our environment, but between ourselves and others, potentially affecting interpersonal exchanges (Heed et al., 2010; Teneggi et al., 2013; Patanè et al., 2017). Indeed, a growing body of evidence suggests that PPS is influenced by the mere presence of others and the nature of interactions with them. Studies in monkeys (Ishida et al., 2010) and humans (Brozzoli et al., 2013) reveal that premotor and parietal neurons respond to visual stimuli not only in one's own PPS but also in others' PPS. For example, Ishida et al. (2010) discovered visuo-tactile neurons in the macaque's VIP area that

responded to tactile stimulation on the animal's body and to visual stimuli approaching both the animal and an experimenter facing it. Similarly, Brozzoli et al. (2013) showed that in humans, vPMc regions were modulated not just by the distance of an object from the participant's hand, but also by the distance of that object from another person's hand, showing mirror-like mechanisms. Teramoto (2018) found that delivering a tactile stimulus on one's hand, the participants exhibited shorter detection times when the disk approached the hand in their near space. When performing the task with a partner across the table, the participants exhibited shorter detection times for the tactile stimulus both when the disk approached their own hand in their near space and when it approached the partner's hand in the partner's near space but the participants' far space. This phenomenon was also observed when the body parts from which the visual stimuli approached/receded differed between the participant and partner (Teramoto, 2018). Research on VTI effects has demonstrated that viewing someone else's face being touched facilitates one's tactile detection of their own face (Serino et al., 2008; Làdavas & Serino, 2010). Thomas et al. (2006) demonstrated that detecting a tactile stimulus on one's own body can be influenced by prior visual cues on a model's body, either enhancing or delaying the tactile identification depending on the visual cue location congruency according to CCE rules.

These findings all indicate that the mechanisms involved in the integration of multisensory stimuli can respond similarly to stimuli in one's own PPS and those of others, i.e., these systems are sometimes said to 'map the space around others'. This ability to integrate visuo-tactile stimuli in another person's space can also affect the representation of one's own. For instance, Heed et al. (2010) found that visuo-tactile interaction strength, measured by CCE, decreases when another person enters the participant's PPS and performs a similar task. Further evidence from dynamic audio-tactile interaction tasks suggests that the size of an individual's PPS can be altered by the mere presence of others. Teneggi et al. (2013) found that looming sounds affected tactile processing at closer distances when participants faced a stranger, compared to when they faced a mannequin, implying a contraction of PPS in the presence of an unknown individual. However, the author found that subsequent social interactions shaped PPS representation: following a fair interaction, participants' PPS expanded to include the space around the other person, whereas no such expansion occurred after an unfair interaction (Teneggi et al., 2013). This finding suggests that after positive social exchanges, PPS might extend to create a shared interaction space, even when there is no physical exchange.

To investigate such social regulation, some recent studies have focused on variations in interpersonal distance as an effect of social cognition (Iachini et al., 2014, 2015; Patané et al.,

2017), indicating that emotional and social factors likely modulate how PPS neurons map the space around others into one's own (Cardini et al., 2012). Evidence suggests that attributes of others - like their moral character, emotional state, and other complex cognitive factors—significantly influence our spatial representation and sensorimotor mechanisms, shaping how we perceive and respond to others in shared spaces. These findings suggest that PPS is responsive to physical proximity, but also to social and emotional cues from others. This has led to a new theoretical perspective in which PPS is seen as a dynamic interface that facilitates interaction between the self and objects, but also other people (Blanke et al., 2015; Cléry & Hamed, 2018; Serino, 2019).

Furthermore, Patanè and colleagues (2021) explored how the sense of ownership attributed to an object that is the target of one's movement, modulates the extent of visuo-tactile interference, in an interpersonal multisensory motor interaction. In this experiment, conducted in pairs, participants were asked to detect tactile stimuli while ignoring visual cues appearing on an object they were about to grasp, using a setup similar to the cross-modal congruency task developed by Brozzoli and colleagues (2009, 2010). Participants performed reach-to-grasp movements toward a glass-shaped object, which came in two distinct versions: each participant was assigned a different glass, marked with a unique coloured band around the rim, to establish a sense of ownership. The task involved various conditions: participants either reached for their own object, reached for their partner's object, or observed their partner performing the grasping action. This design allowed the researchers to examine how the sense of ownership, and the interaction between participants, influenced the integration of visual and tactile information during the action. What they found points in favour of an ownership-dependent effect, VTI increased when participants interacted with their own object or observed their partner grasping their own object. However, this effect was absent when participants interacted with or observed interactions involving an object they did not own. When both participants shared a single glass-shaped object, the VTI was observed as for the case of ownership. This suggests that the modulation of PPS representation may depend on the sense of ownership over the object being interacted with. Importantly, these findings highlight that a social component can influence how the brain integrates sensory information in interactive contexts.

However, despite these results on the plastic 'extension' of the multisensory self-space representation in social contexts, PPS still serves a function in distinguishing the self from the environment and others, establishing a stable sense of bodily self-awareness (Blanke, 2012; Ehrsson, 2012; Makin et al., 2008; Tsakiris, 2010). Indeed, PPS is believed to contribute directly to the experience of self-location, the subjective sense that one occupies a particular position in space

(Blanke et al., 2015; Cléry & Hamed, 2018; Legrand et al., 2007; Serino, 2019). In line with this framework, but in contrast to previously cited works, Fossataro et al. (2023) demonstrated that the proximity of another person's hand can shrink the area of space where multisensory stimuli are processed, hence the boundaries of PPS. This suggests that the PPS model, based on bodily and sensory input, adjusts dynamically to account for the presence of others, redefining the boundaries between self and other. When space is shared, this plasticity may help distinguish external events as relevant to the self or to others.

In general, the evidence reviewed above shows that there are many different features of social contexts (e.g., the mere presence of another, their emotional or moral connotation, whether we have interacted with them, whether an object belongs to another) that may affect one's PPS representation (as indexed by multisensory integration effects) and consequently, interpersonal interactions. Thus, a comprehensive understanding of how specific social factors in interpersonal interactions influence the sensory-motor representation of events surrounding one's own body and that of others remains an ongoing area of research.

## **1.5. Social Factors Influence on Interpersonal Interactions**

One's movement play a crucial role in para-verbal communication (Pollick & De Waal, 2007; Arbib et al., 2008; Xu et al., 2009; Corballis, 2010; Vainiger et al., 2014): the use of gestures and signals help align individuals, enabling effective coordination through signalling-and-decoding strategies (Candidi et al., 2015). However, beyond these explicit social signals, other characteristics of the partners—such as their physical aspects (i.e., age, gender, height; Uzzell & Horne, 2006; Iachini et al., 2016; Hecht et al., 2019), moral attributes (Iachini et al., 2015; Pellencin et al., 2017; Fini et al., 2020), group affiliation (Willis, 1966; Leibman, 1970; Tajfel et al., 1971; Fini et al., 2020) and the emotional valence of their aspect or their actions (i.e. degree of threat of the social stimuli; Lockard et al., 1977; Hayduk, 1983; Siegman & Feldstein, 2014; Ruggiero et al., 2017; Cartaud et al., 2018, 2020) —also critically influence how individuals adjust their behavior when interacting or communicating with each-other. The following sections critically review the available literature on the interaction between interpersonal perception in action, spatial and personological sense (i.e., similarity/distance in personality traits/states) and linguistic (non-verbal but written) communication.

### **1.5.1. From Sensorimotor Contingencies to Social Cognition**

The close relationship between cognition and action, through which individuals experience and interpret the world, has led pragmatist theories to conceptualize cognition as inherently embodied within action. (Varela et al., 1991; Clark, 1997; Noë, 2004; Pfeifer & Bongard, 2006; Engel, 2010). In line with this view, the concept of Sensorimotor Contingencies (SMCs) states that cognition can be understood only in association with movement and sensory inputs, which are entangled in a bidirectional relationship (O'Regan & Noë, 2001). SMCs offer a conceptual framework for exploring how individuals process and formulate object-related concepts and action strategies. The successful acquisition of SMCs enables goal-directed actions, expanding their significance beyond basic movement-sensory input to cognition (Maye & Engel, 2011, 2012; Engel et al., 2013 ; Högman et al., 2013 ). This perspective goes beyond the original sensorimotor contingency theory (SMCT), which emphasized immediate sensory feedback from movements (O'Regan & Noë, 2001). Instead, it supports the idea that intricate interactions arise through the continuous dynamic pairing between individuals and their environments.

In their environments, individuals are involved with other conspecifics. Hence SMCs has also been proposed to have a social dimension (socSMCs), which functions as an action-centered mechanism for social cognition (Lubbert et al., 2021). The concept of socSMCs highlights how basic sensorimotor interactions are fundamental to foster social entrainment and engagement. Conversely, social contexts play a significant role in shaping these low-level sensorimotor relations. For instance, shifts in social behaviors, such as variations in proximal or distal sensorimotor coupling, can impact higher-order social cognitive processes like trust, empathy, and cooperation (Froese et al., 2014; Keller et al., 2014; Llobera et al., 2016). Likewise, changes in social involvement may alter the intrinsic and extrinsic dynamics of sensorimotor coordination between two individuals (Lubbert et al., 2021; Sacheli et al., 2015b). Indeed, simulative mechanisms associated with the fronto-parietal mirror system (Rizzolatti & Craighero, 2004) relates to basic neurophysiological responses to interpersonal interactions, which are influenced by higher-order cognitive and cultural factors. Neuroimaging research has pinpointed specific brain regions and networks that become active through mentalization, empathy, and mirroring tasks (Stanley & Adolphs, 2013). Furthermore, the connection between motor control and social cognition is highlighted by studies on motor mimicry, an automatic unconscious form of interpersonal coordination enabled by the mirror neuron system (Wang & Hamilton, 2012). Additionally, empathy is found to arise from fundamental sensorimotor processes and shared representations

involved in both executing and observing actions (De Waal & Preston, 2017). Collectively, these studies emphasize the intricate interplay between motor and social dynamics in interpersonal exchanges.

Theoretical frameworks on self-representation provide essential insights into how individuals perceive themselves and interpret social interactions, since self-identity formation is intrinsically linked to social and sensorimotor processing. Illustrating how self-perception and group dynamics influence interpersonal behaviour, Tajfel (1971, 1982) highlighted the significant influence of in-group versus out-group distinctions on self-perception. Individual differences and cultural influences significantly impact neural and behavioural responses within social environments (Kubota et al., 2012). How we perceive ourselves is fundamental in shaping our relationships with others. Group identification is crucial for fostering self-esteem and self-concept, often resulting in favouritism toward one's own group while occasionally leading to the denigration of out-group members (Chang et al., 2014; Mackie & Smith, 2002). Farmer and Maister (2017) distinguish between two types of self-representation: the bodily self, which is grounded in sensorimotor processing, and the conceptual self, which comprises associative and semantic knowledge developed through social interactions. Both aspects of self-representation are connected with social cognition (Farmer & Tsakiris, 2012) and play a role in forming and perpetuating social biases (Farmer & Maister, 2017).

Racial biases, shaped by the skin colour of another individual, have been shown to alter sensorimotor mirroring processes when individuals observe both neutral actions and emotional expressions (Molnar-Szakacs et al., 2007; Desy & Théoret, 2007; Gutsell & Inzlicht, 2010). Group membership also plays a crucial role in social attention, which can be assessed through gaze-mediated orientation (Pavan et al., 2011). Furthermore, these biases extend to the empathic mapping of sensorimotor responses and emotional reactions toward others experiencing pain (Avenanti et al., 2010; Azevedo et al., 2013). Evidence from neuroimaging studies highlights the critical roles of the medial prefrontal cortex (mPFC) and the anterior insula (AI) in the neural basis of prejudice (Amodio, 2014). These regions are particularly involved in processing information regarding individuals' traits, preferences, and beliefs (Frith & Frith, 2003; Ma et al., 2013; Schilbach, 2015). Additionally, these regions display sensitivity to various social factors, including social status and experiences of ostracism (Muscatell et al., 2012; Powers et al., 2013). Moreover, ethnic categorization significantly influences neural activation within the mirror neuron system, particularly during the comprehension of intentions (Liew et al., 2011) and the act of imitation (Losin et al., 2012; Earls et al., 2013), meaning that these processes engage neural networks

differently when individuals are confronted with their own ethnic group versus other groups. Interestingly, prejudice appears to be implicitly modulated through embodied experiences, particularly through bodily illusions (Farmer & Tsakiris, 2012; Maister et al., 2013; Peck et al., 2013). Moreover, research has demonstrated that multisensory stimulation can influence lower-level sensorimotor indicators of prejudice (Fini et al., 2013; Maister et al., 2015; Hasler et al., 2017).

In conclusion, these findings indicate that our self-representations, bodily and cognitive experiences can subtly shape our social perceptions and that social cognition can actively influence our interaction dynamics.

### **1.5.2. From Communicative Gestures to Language**

Communicative gestures act as precursors to language acquisition, sharing their neural basis on both ontogenetic (Petitto & Marentette, 1991; Iverson & Goldin-Meadow, 2005) and phylogenetic levels (Hewes et al., 1973; Pollick & De Waal, 2007; Arbib et al., 2008). Beyond merely conveying information, some gestures invite specific responses from the observer, functioning as pre-verbal social significant signals that call for an action. Neuroimaging and lesion studies provide evidence indicating that the perception of objects can prime motor responses, namely affordances, even without a conscious intention to interact with those objects. This process is governed by a left-lateralized fronto-parietal network responsible for preparing motor actions (Grafton et al., 1997; Grèzes et al., 2003; di Pellegrino et al., 2005; Buccino et al., 2009). In line with a non-externalist view of affordances (Chemero, 2003; Caiani, 2014; Borghi, 2018a), they are not just properties of objects but are deeply embedded in sensorimotor patterns that invite for action.

Extending this concept to social contexts, social affordances are opportunities for interaction embedded in gestures and other social cues (Valentini & Gold, 1991; Constall, 1995). In this view, interactive gestures, are linked to the perception of others' social responsiveness. This concept implies that gestures can entail social affordances. In interpersonal contexts, interactive gestures are observed to trigger sensorimotor responses in the observer to complete a joint action (Curioni et al., 2020). This system is key in modulating our readiness to act based on social cues. Rietveld and Kiverstein (2014) further extended the notion of affordances by associating them with the skill set of organisms, particularly humans, where affordances are tightly linked to sociocultural practices (Wokke et al., 2016; Iacoboni et al., 2005). This perspective suggests that perception involves an

‘openness to affordances’ (Rietveld & Kiverstein, 2014), meaning that individuals are attuned to their environment's social and cultural cues.

A growing body of evidence suggests that linguistic context significantly impacts affordance processing. For instance, labelling a novel object can facilitate motor learning related to using that object, such as grasping it, while being less effective for tasks like moving it (Lupyan, 2012). Language encodes stable sensorimotor features, such as shape and size of objects, which are relatively context-invariant, while more dynamic features like object orientation may be less consistently encoded. This points to language’s essential role in grounding conceptual knowledge and facilitating motor learning (Foerster et al., 2020). Furthermore, specific verbs can differentially activate motor affordances (Borghi & Riggio, 2009) and merely listening to action-related sentences can prime motor responses (Gianelli et al., 2013). Social variables, such as familiarity with the agent, also modulates affordance activation (Gianelli et al., 2013), highlighting the combined influence of language and social context.

According to the Distributed Domain-Specific Hypothesis (DDSH; Caramazza & Shelton, 1998) the brain organizes knowledge based on evolutionary relevance, with category-specific knowledge stored together. Particularly, concrete word meanings are tied to specific neural regions that process perceptual features such as color and taste (Simmons et al., 2007; Barrós-Loscertales et al., 2012). Unlike concrete concepts, which are learned through direct physical interaction (Schwanenflugel, 1991; Paivio, 1986; Borghi et al., 2017), abstract concepts are largely acquired through language in interpersonal interactions (Paivio, 1991; Barsalou, 1999; Brysbaert et al., 2014). Embodiment theories suggest that sensorimotor activation is essential for semantic processing, alternative views propose that such activation may be secondary to lexical processing (Mahon & Caramazza, 2008). Hence, language is essential in acquiring and processing abstract concepts (Bergelson & Swingley, 2013; Borghi & Setti, 2017; Lewis et al., 2021), as shown by studies linking abstract concepts to linguistic systems more than concrete ones (Wauters et al., 2003). However, despite abstract concepts being traditionally viewed as detached from sensorimotor systems, they are increasingly recognized as being influenced by them. For example, different categories of abstract concepts, such as mathematical and emotional, involve distinct brain regions, often related to motor or sensory functions (Dehaene, 1997; Desai et al., 2018). Concepts such as ownership can affect motor behavior and spatial perception (Constable et al., 2011; Borghi et al., 2018a; Patanè et al., 2021).

The Words as Social Tools (WAT) theory posits that abstract concept acquisition is primarily driven by linguistic and social interactions, with inner speech playing a crucial role (Borghi et al., 2019 a, b; Fini et al., 2022). These interactions necessitate coordinated communication,

underscoring the idea that language itself is a form of joint action (Clark, 1996). This aspect is especially evident in discussions on concepts such as freedom, which are flexible and context-dependent, requiring complex cognitive and social negotiation during communication (Borghi et al., 2019a,b). Conversations about abstract concepts, particularly in socially or politically charged contexts, can serve as mechanisms for reinforcing social bonds and fostering group cohesion according to cognitive evolution theories (Rietveld & Kiverstein, 2014; Wokke et al., 2016). Thus, abstract concepts serve as cognitive tools and function as social instruments, facilitating a shared understanding (Borghi et al., 2019a,b). Consequently, discussing these concepts could promote psychological and physical proximity among individuals and shape interpersonal interactions.

## **1.6. Thesis Contributions and Hypotheses**

The present work aims to deepen our understanding of how interpersonal sensorimotor processes relate with multisensory integration and social cognition in interactive contexts, at both behavioural and neurophysiological levels. While the second chapter focuses on how sensorimotor processes, such as VTI effects and motor control, adapt to interpersonal interactions within shared interpersonal spaces, the third chapter explores how higher-level social factors, such as language, influence interpersonal sensorimotor and cognitive dynamics.

Specifically, previous research showed that visuo-tactile interference expands when planning to act on distant objects (Brozzoli et al., 2009, 2010; Patanè et al., 2019; Senna et al., 2019). Here, in Chapter 2, across three experiments, VTI are tested in a joint action task with a virtual partner in synchronously grasping an object together (Chapter 2.1.). The research addresses how multisensory integration dynamically adapts to the partner's movements, reshaping sensory representations within shared interaction spaces. In line with the concept of 'Socializing Sensorimotor Contingencies' (socSMC), integrating sensorimotor signals from interacting partners can lead to the emergence of shared sensorimotor representations (Lubbert et al., 2021), behavioural and neural patterns evoked by visuo-tactile stimuli might be influenced on prior joint interactions. Therefore, Chapter 2.2. investigates how prior interpersonal interactions involving different sensory channels influence subsequent sensory processing responses to cross-modal stimuli. Furthermore, previous studies have demonstrated how synchronous visuo-tactile stimuli can alter the perception of one's body ownership (Botvinick & Cohen, 1998; Tsakiris & Haggard, 2005). Chapter 2.3 explores the phenomenon of body ownership and multisensory integration within a novel 'social' rubber hand illusion (RHI) paradigm. This study introduces an interpersonal dynamic to the classical RHI,

where participants engage in synchronous mutual stroking with a partner. The study investigates how this reciprocal interaction modulates not only body ownership but also subsequent VTI, expanding on earlier work on self-other body representation (Blanke et al., 2015; Cléry & Hamed, 2018; Serino, 2019; Fossataro et al., 2023).

Chapter 2.4. delves deeper into the involved neural mechanisms of social cognition in interpersonal interactions, particularly in the context of racial biases. Studies suggest that the medial prefrontal cortex (mPFC) plays a key role in social biases, as it is involved in regulating racial stereotyping and empathy (Amodio, 2014; Decety et al., 2009; Stürmer et al., 2006). Hence, Chapter 2.4., explores how racial group membership can impact sensorimotor processes, influencing coordination during joint actions with out-group members. In this study, the role of the mPFC in joint actions with out-group partners was explored using inhibitory continuous theta burst stimulation (cTBS). This research expands previous findings showing that out-group membership reduces and disrupts predictive simulation of a partner's movements (Sacheli et al., 2015b).

As social interactions happen in specific social contexts, Chapter 3 expands on investigating how social factors, such as conversational topics, interact with and modulate interpersonal dynamics at the behavioural and subjective level. Specifically, Chapter 3.1., highlights the role of language in shaping the quality of social interactions. Since abstract concepts (e.g. democracy, justice) are primarily acquired through linguistic interactions rather than sensory experiences, as for concrete concepts (Wauters et al., 2003), their meanings are deeply rooted in social interactions (Borghi, 2022; Fini et al., 2021, 2023a). How abstract versus concrete words influence dialogue initiation and maintenance is an open question addressed. Abstract language encourages participants to reflect and explore broader topics, which can break the ice in initiating dialogues. This study investigates whether abstracts, compared to concrete concepts, promote interpersonal connection by enabling the negotiation and establishment of shared meanings essential for effective communication (Borghi & Tummolini, 2020; Gilead et al., 2020). In line with the notion that conversational exchanges are a form of joint activity, where effective collaboration leads to better communication (Clark, 1996; Garrod & Pickering, 2004), Chapter 3.2., explore how brief online chats can significantly boost interpersonal closeness, regardless of the conversational topic (abstract/concrete). The study involved participants engaging in short virtual conversations and measured their sense of connection afterward, expanding on the idea that computer-mediated communication can still foster meaningful connections. This study highlights the potential for digital platforms to enhance interpersonal relationships. Building on this broader understanding of how social interactions foster interpersonal closeness but are also influenced by social cognition,

This work contributes significantly to the ongoing exploration of several open questions in the fields of multisensory integration and social cognition in interpersonal interaction. Overall, the present thesis contributes to our understanding of how sensory-motor processing and social context influence interpersonal interactions, focusing on the dynamic nature of individuals' exchanges. Exploring how individuals coordinate and interact within shared spaces influencing multisensory integration (e.g., visuo-tactile interference), motor control (e.g., action coordination) and interpersonal representations (e.g., after language use). This integrated approach sheds light on the mechanisms underlying our ability to adapt to and navigate social environments, offering insights into the dynamic nature of human interactions.

**2**

**Unraveling Multisensory Integration in  
Interactive Contexts**

## **2.1. On-line Interpersonal Motor Interactions Modulate Visuo-Tactile Integration Mechanisms**

### **2.1.1. Introduction**

Events that happen close to the body are extremely important because their efficient perception and representation guide goal directed (Graziano & Gross, 1995; Maravita et al., 2003a,b) and defensive behaviours (Cooke & Graziano, 2003). The representation of events and objects in the region of space immediately surrounding the body, called the peri-personal space (PPS; Rizzolatti et al., 1981a, b; Brozzoli et al., 2011), is supported by a number of multisensory mechanisms that generate sensory facilitation (Spence et al., 2004), reduction of ambiguity (e.g. by shaping body and space representation, Serino et al., 2015), perceptual (Maravita et al., 2003a) and action accuracy (Brozzoli et al., 2010). A classic example of these intersensory effects is the cross modal congruency effect (CCE; Spence et al., 1998, 2004) where a visual (Pavani et al., 2000; Spence et al. 2000) or auditory (Serino et al., 2007, 2011) cue happening close to a body part, or approaching it (Graziano et al., 1997; Canzoneri et al., 2012; Clery et al., 2015; Clery et al., 2017), facilitates the detection of a tactile stimulus delivered on the corresponding body part (Farnè et al., 2005; Brozzoli et al., 2012).

There is extensive evidence that multimodal fronto-parietal neurons (Fogassi et al., 1992, 1996, 1999) and neurons in the putamen (Graziano and Gross 1993, 1995; Graziano et al., 1994) integrate visuo-tactile stimuli close to the body and underpin these effects. These neural systems support the behavioural evidence of somatotopically organised peri-hand (Brozzoli et al., 2012), peri-trunk (Serino et al., 2015) and peri-face (Ishida et al., 2010; Farnè et al., 2005a) spaces. Furthermore, training to use a rack to retrieve food (Iriki et al., 1996) or perform actions (Maravita et al., 2003a, b) may plastically expand the visual receptive field of visuo-tactile neurons and concurrently expand the PPS (Farnè et al., 2007).

Recent computational models of PPS (e.g., Magosso et al., 2010a, 2010b; Serino et al., 2015) propose that tactile stimuli on the body and visual or auditory stimuli near the body, but not far from it, are integrated. These models include unisensory layers for tactile, visual, and auditory stimuli, which project to multisensory areas with strong synapses tuned to respond preferentially to near-body stimuli. The synaptic weights of the receptive fields for far-space unisensory visual or

auditory stimuli are weaker, thus limiting activation of multisensory neurons to stimuli occurring close to the body (Magosso et al., 2010a; Serino et al., 2015). This computational framework explains in part how multisensory integration in the PPS can support the brain's ability to represent near-body space dynamically through these mechanisms of reference frame transformations, which re-code sensory inputs into body-centered coordinates (Makin et al., 2013; Noel et al., 2018).

More recently it has been shown that visual cues outside the PPS can still facilitate tactile processing in case the visual cue appears far from the body but at the location of an intended action or during its execution (Brozzoli et al., 2009; Noel et al., 2015; Senna et al., 2019; Patané et al., 2019). This cross modal congruency effect associated with movement execution has been described also when participants observe other people grasping objects that are associated with them (Patané et al., 2021). These recent behavioural results open the possibility to imagine that when performing movements with a partner we need to be extremely accurate in predicting the sensory (visual, auditory and tactile) consequences of their movements and integrate them with the sensory consequences of our own movements. While the mere presence of another individual in front of us is able to “expand” (Serino, 2019) and “contract” the extension of the PPS (Fossataro et al., 2023), by modulating perception of tactile stimuli over our body (Heed et al., 2010; Thomas et al., 2006; Fossataro et al., 2023), others have suggested that interacting with others might create a Shared Action Space (Pezzulo et al., 2013). One possibility is that multisensory events happening in the space of the interpersonal interaction are processed, and integrated, following different compatibility rules compared to those at play when being static or moving alone or in front of another individual but being still.

Here, in three behavioural experiments, we tested the hypothesis that, when on-line coordinating with others by performing same or opposite reaching and grasping movements, the cross modal visuo-tactile integration (VTI) effect found during grasping execution (Brozzoli et al., 2009; Patané et al., 2019, 2021; Senna et al., 2019), i.e., the facilitation of tactile processing at the fingertip associated to the presentation of visual cues appearing on an object at the location of a reaching and grasping movement, is remapped according to the movement of the interaction partner, not one's own. Experiment 1 investigates VTI modulations according to the rules of interpersonal interactions: we expect that multisensory effects during a joint grasping task are influenced by both the participant and the partner's movements as a function of the interaction type. Specifically, when coordinating in a reach-to-grasp movement, complementary interaction would improve agent's detection of incongruent visuo-tactile stimuli compared to an imitative interaction. Experiment 2 is directed to examine whether the results from the first experiment would be replicated and to test

that the effect is intersensory in nature, rather than driven by visual dominance. Visual dominance refers to the tendency for visual information to take precedence over information from other sensory modalities when they are presented simultaneously (Rock & Victor, 1964; Spence & Driver, 2004). If so, the perception of tactile stimuli might be altered by the presence of simultaneous visual stimuli, leading to a stronger influence of visual information on the overall sensory integration process. To test this, we introduced a unisensorial visual condition, in which a visual cue would be presented in absence of any tactile stimulation. If the results from Experiment 1 are due to a multisensory integration of the visual and tactile stimuli, we should expect a similar PPS recruitment when visual congruent or incongruent stimuli are presented, but also when visual and tactile information are combined. Experiment 3 aimed to test the interactive nature of the effect observed in the previous two studies by introducing an anticipated visuo-tactile stimulus presentation, delivered when participants couldn't predict the avatar's directional goal yet. If the effect is modulated by the interacting rules, the performance (measure in reaction times, RTs) in detecting tactile stimuli during incongruent trials should be contingent on the predictability of the avatar's reach-to-grasp end location (not in the anticipated visuo-tactile stimuli presentation). Therefore, when the visuo-tactile stimulus would be presented too early to anticipate the partner's movement, we would expect no modulations of VTI. Collectively, these findings could provide new insights into the flexibility of low-level sensory mechanisms in supporting highly complex tasks such as interpersonal interactions.

### **2.1.2. Materials and Methods**

This section comprehensively outlines the shared methods employed across the three distinct experiments, providing an overview of the procedures adopted. In all experiments participants were involved in an interpersonal motor task while receiving tactile stimuli over their finger(s). They were asked to synchronise their reach and grasp movements with those of a virtual partner in different conditions while their main task was to detect as soon as possible the tactile stimulus.

#### **Common Set-up**

##### **Interpersonal Reaching and Grasping task**

We adapted a well-established human-avatar interactive task, that combined an ecological set-up with experimental control (Sacheli et al., 2015a; 2015b; 2018; Candidi et al., 2017; Era et al., 2018;

2020; 2023; Gandolfo et al., 2019; Fini et al., 2021). This task had previously demonstrated its ability to engage the same behavioural mechanisms observed in human-human interactions, specifically mutual adjustment and automatic imitation (Sacheli et al., 2012; 2013; Candidi et al., 2015; Curioni et al., 2017; Era et al., 2018; Boukarras et al., 2021). Participants were seated in front of a rectangular table and an LCD monitor was positioned at the opposite edge of the table, 60 cm away from participants' eyes where the interaction partner was shown (see below). The task assigned to participants required them to reach and grasp a bottle-shaped object (37 cm in height) positioned next to the middle of the table, between them and the monitor and to touch the object in synchrony with their partner (Figure 2, 6, 9). Participants were required to perform a reaching-to-grasp movement in synchrony with an avatar's movement. Considering the diameter of the bottle-shaped object, made of two rectangles in a stack with distinct widths, they could either perform a precision grip (on the top, 2.7 cm diameter) or a power grip (on the bottom, 6.5 cm diameter) with their right hand. Prior to each trial, participants were instructed to press and hold, with their index finger and thumb, a start button located on the right side of their midline (at 34 cm distance from the bottle object and 10cm away from the midline). Following an auditory signal played through headphones, they released the button to initiate the grasping movement together with the virtual avatar presented on the screen in front of them. Grasping times were measured using touch-sensitive markers placed in pairs on each rectangle, precisely at distances of 15 cm and 22 cm along the vertical height of the object.

Two different auditory stimuli were used to instruct participants in the task, depending on the experimental mode of interaction. Indeed, the experimental task involved two conditions that differ in the interactivity level (Interaction Mode Factor): (1) in the Interactive condition, the goal of participants' action cannot be achieved without considering that of the virtual partner, taking into account its movements and adapting to them in real time; (2) in the Cued condition, participants do not need to coordinate their behaviour to that of the partner, while grasping in synchrony with it. One auditory stimulus was associated to the Interactive condition in which participants were instructed to perform complementary (e.g., if the virtual partner grasped the lower bottle part, the participant grasped the upper part, and vice versa) or imitative (e.g., both grasping the upper or lower part of their bottles) movements in relation to the avatar. Before the commencement of each trial, participants received audio instructions in the form of vocal commands, specifying whether they had to perform same or opposite movements. Conversely, the auditory stimulus associated with the Cued condition, instructed participants that they would hear one of two tones indicating whether they should grasp the upper or lower part of the bottle. A high-pitched tone (1470 Hz)

signalling to reach toward the upper white panels, while a lower-pitched tone (114 Hz) indicated reaching toward the lower white panels. Each of the four auditory cues had a duration of approximately 193 ms. Prior to initiating different conditions, participants were instructed on the corresponding auditory cues. To ensure participants comprehension of the instructions, a set of try-outs was conducted before each condition. Once the trial runs were successfully completed, the actual experiment commenced and the starting condition was randomised and counterbalanced between participants. In both Cued/Interactive conditions, the final participants' goal was to grasp the bottle-shaped object as synchronously as possible with the virtual partner. For both interaction modes, the movements, with respect to the partner's grip, could either be Same or Opposite (Movement Type Factor); depending on whether they grasped the bottle with the same grip or in a complementary manner relative to the avatar.

The trial timeline was as follows: participants were presented with either the Interactive (Same/Opposite) or Cued (Up/Down) auditory instruction via headphones, followed by a fixation cross lasting for 300ms. Afterward, the Avatar's appearance on the screen (see Visual Stimuli of the Avatar Reaching and Grasping Movements section) marked the soon beginning of its movement, which occurred between 200ms and 600ms. Upon receiving the auditory instruction and seeing the avatar, participants were allowed to release the start button and synchronously move to grasp the bottle-shaped object along with the avatar. The Avatar executed its action during a trial duration of approximately 2000ms. The time between trials was determined by participants' return to the starting position, with the subsequent trial initiated by the experimenter upon their return on the "start" button. Trials where participants initiated their movement prior to the auditory instruction were excluded from the analysis. A valid trial denoted participants' adherence to the auditory instructions, accurate execution of the bottle grasp and detection of the tactile stimulation presented. Presentation and randomization of the stimuli was administered through E-Prime2 Professional software (Psychology Software Tools Inc.).

### **Visual Stimuli of the Avatar Reaching and Grasping Movements**

As in previous studies (Sacheli et al., 2015; Candidi et al., 2017; Gandolfo et al., 2019; Moreau et al., 2020; Era et al., 2020; 2023), the kinematic attributes of the virtual partner were derived from the movements of human participants engaged in different grasping actions with the right dominant hand within a joint-grasping task involving human-human interactions (refer to Tieri et al., 2015; Sacheli et al., 2015; Fusaro et al., 2019; for comprehensive technical information about Motion Capture recording). Motion capture was performed using 10 infrared light-emitting cameras and the

3D positions of 37 passive reflecting markers (5 mm diameter), attached to the participant's entire upper body (pelvis, chest, head, left and right arm, and right hand), recorded with a spatial error below 1.5 mm and at a temporal resolution of 120 Hz. Raw data were processed offline using commercial Vicon MX optical tracking software (Vicon Motion Systems, Oxford, UK) to reconstruct and label the markers and to interpolate short missing parts of the trajectories. Utilising MotionBuilder 2017 and 3DS Max 2017 (Autodesk) software, the processed trajectories were accurately generated and applied to a male character of Caucasian ethnicity. Hence, the stimuli represented a male avatar conducting grasping movements toward a bottle-shaped object, as the one presented to the participant in the experimental set-up.

The avatar featured solely the upper body, spanning from the shoulders downward, excluding the neck and head intentionally to avoid diverting subjects' attention to the virtual partner's facial expressions. A total of 10 distinct grasping motions were captured in the video clips dataset. Half of these motions ended with the avatar's hand gripping the lower section of the bottle-shaped object (power grip), while the remaining half concluded with the avatar's hand grasping the upper portion (precision grip). An additional 4 clips were crafted, depicting digital avatars executing movement adjustments by transitioning from precision to power grip (or vice versa) during the reaching phase. These modifications were applied in 20% of the trials wherein online corrections were involved. The process of crafting correction videos involved combining the initial frames of one clip, like a power grasp clip, with the final frames of another, such as a precision grasp clip, accomplished using 3DS Max. These modified clips were considered as catch trials, so not relevant for the analysis, but to keep the participants' engaged in the task. Thus, the 14 clips could be divided into four conditions (correction/no-correction  $\times$  power/precision grip): 5 precision grips, 5 power grips, and 4 corrected trajectories, two per grip. These clips were pre-recorded rather than generated dynamically but were randomized trial-by-trial to ensure variability within the task. While this approach maintained an element of variability across trials, it is true that the finite number of clips may limit the range of naturalistic variability compared to real-life movements. Additionally, all video stimuli had a duration of approximately 2000 ms, altering the duration of movements between 1600 ms and 2400 ms to better align with participants' behavior and avoid potential habituation effects. In previous research, a pilot study was conducted to validate the naturalness of the avatar's movements (Sacheli et al., 2015). Twelve participants rated the avatar's movements as realistic and natural ( $63.8 \pm 24.4$  on a visual analogue scales ranging from 0 to 100). Moreover, a two-way repeated measure ANOVA on the different types of clip (correction/no-correction trajectory  $\times$  precision/power grip) did not show any significant difference (all p values  $>0.2$ ). These findings

suggest that the approach was effective in preserving ecological validity while ensuring experimental control.

The grasp times of the avatar were logged on a trial-by-trial basis using a photodiode situated on the screen. These two photodiodes placed on the screen transmitted their signals to the tactile stimulator and to a computer that was then recorded their timing by E-Prime 2 Professional (version 2.0.10.242, Psychology Software Tools Inc., Pittsburgh, PA) through a TriggerStation (BrainTrends Ltd., Italy). The activation of one photodiode was initiated by a white square positioned at the bottom left of the screen (unseen by the participants) to trigger the tactile stimulation. Additionally, a second photodiode was activated by another white square placed at the bottom right of the screen (not visible to the participants), which coincided with the moment when the avatar virtually grasped its object.

### **Visuo-Tactile stimulation and task**

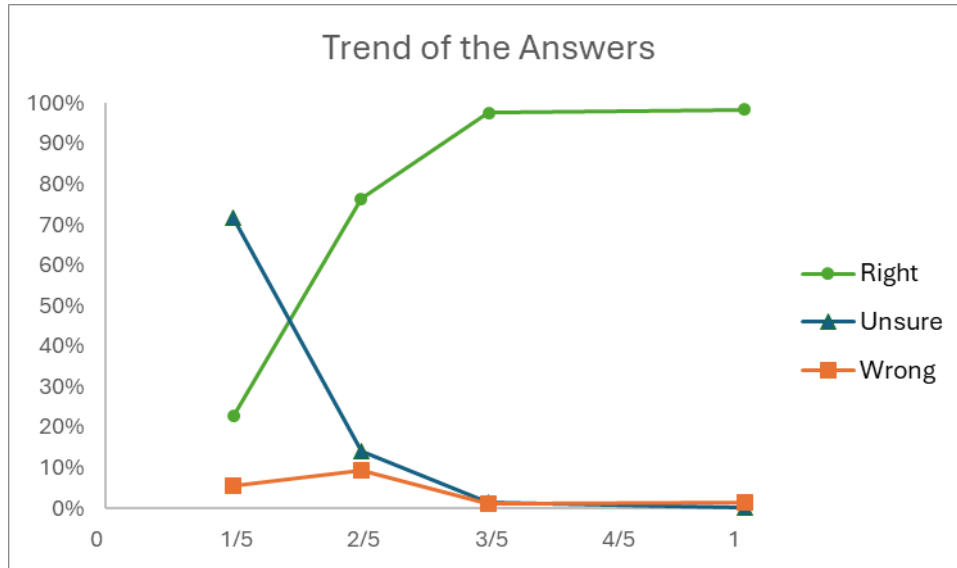
While performing the grasping task, participants were instructed to react to a tactile stimulation. Tactile stimulation was administered to participants using a constant-current electrical stimulator (DSA7A, Digitimer Ltd., Welwyn Garden City, United Kingdom). Disposable electrodes were positioned dorsally on both the index and the thumb (Experiment 1) or the index finger (Experiment 2, 3) of the participants' right hand with which they performed the prehension movement (see Interpersonal Reaching and Grasping Task section), maintaining a 1 cm separation between them. To ascertain stimulus intensity, a tactile threshold procedure (Serino et al., 2011) was employed. Starting with minimal levels of tactile stimulation (100-400 Vmax), the intensity was incrementally increased until participants reported a clear sensation. Following the initial reported level of tactile perception, participants were stimulated 10 times and requested to report whether they felt the stimulation. If participants successfully detected it 90% of the time (9 out of 10 instances), that value was adopted as the tactile stimulation threshold for the rest of the experiment. The duration of the tactile stimulus was 200 $\mu$ s. During the experiments, participants were asked to detect the tactile stimulus by pressing a keyboard/pedal response.

Synchronously to the tactile stimulation, a visual cue could appear (80 ms duration) in a position Congruent or Incongruent with respect to the final prehension movement of the participants on their own object (on the top or bottom of the avatar bottle; Visuo-Spatial Congruency Factor). Furthermore, since it appeared on the partner's object, the visuo-tactile stimulation generated another interaction effect with respect to the spatial location of the visual cue and the avatar's grip. In Experiment 1, since both index and thumb fingers were tactile stimulated and the visual cue

could appear in a Congruent or Incongruent position relative to the avatar's final finger position on its bottle (front or back), we measured another interaction effect with respect to the spatial location of the visual cue and the avatar's end-finger position (Visuo-Tactile Congruency Factor). In Experiment 2, the visual cue could be presented with or without the tactile stimulus (Sensory Modality Factor). In Experiment 3 the visuo-tactile stimulation could be presented early or late, relative to the avatar's movement start (Visuo-Tactile Timing Factor).

### **Pilot study**

Prior to conducting the first experiment, a pilot was run, in order to determine the timings within the video clips (see stimuli section) when participants could accurately predict the aimed location of the virtual partner's movement. In this context, video clips were presented to a cohort of 11 healthy individuals (6 Females, mean age  $25.73 \pm 2.11$  SD years), with the clips halted at various stages within the reach-to-grasp action (1/5, 2/5, 3/5, and 5/5 of the total movement duration). Participants were prompted to respond by pressing one of three buttons (uncertain/up/down) to indicate whether, in their opinion, the virtual avatar aimed to grasp the bottle-shaped-object on the upper or lower part or that they were uncertain about the movement's target location. The findings showed that, when the stimulus appeared in the first fifth of the reaching-to-grasp movement, participants were unable to determine the avatar's target direction. In the second fifth of the movement duration, participants started to understand the movement direction, but error rate was still high. Hence, drawing insight from the data (Figure 1), to be sure that participants could accurately discriminate the aimed location of the observed action, visual and tactile stimulation (see stimuli and procedure sections) during the actual experiment has been synchronised at the third fifth of the overall movement duration.



**Figure 1: Trend of answers from 1/5 to 5/5 of the total movement duration for the pilot study.** The chart illustrates the distribution of responses categorised as "Right," "Wrong," and "Unsure" at each time point from 1/5 to the completion of the avatar's movement. Initially, the percentage of correct responses was low (1/5) but steadily increased until reaching a plain (3/5). Incorrect answers showed an upward trend and were most prevalent at the movement's onset, decreasing rapidly at 2/5 and eventually disappearing. Therefore, the 3/5 timepoint seemed to represent the earliest time during the avatar's reach-to-grasp movement at which participants could confidently discern the avatar's movement final location.

## Dependent Variables and General Statistical Analyses

The dependent variables employed for the statistical analysis in the Visuo-tactile Integration (VTI) Task comprises the Reaction Times (RTs) for the Tactile Response (in ms), corresponding to the timing between the appearance of the visuo-tactile stimulation and the discrimination (Experiment 1) or detection (Experiment 2 and 3) response of the participant to the tactile stimulus, measuring VTI effects in the paradigm.

For the Joint Action Task, we considered as dependent variables the Grasping Asynchrony (ms), corresponding to the absolute value of the difference between the grasping time of the avatar and the participant on their own object, measuring the asynchrony of participants' grasp compared to the partner and the overall ability to perform the task.

For both tasks a within-subject Analysis of Variance (ANOVA) was performed, since assumptions for parametric testing, tested through Shapiro-Wilks and Lilliefors test, were generally met.

Accuracy of response (ACC), was considered as the correct discrimination of the tactile stimulation location (thumb/index; Experiment 1) or detection (index; Experiment 2 and 3) of the tactile stimulation and execution of the joint grasping (in all Experiments). We employed nonparametric tests, specifically a Friedman ANOVA, to assess Accuracy.

For all our dependent variables, trials where participants missed the tactile stimulation (i.e., did not initiate keyboard/pedal response) and/or performed anticipatory responses. Furthermore, we excluded trials in which participants failed to grasp the touch-sensitive copper plates (resulting in no recorded response), released the start button before the go instruction, or the complementary/imitative instructions were not followed. Only data from accurate trials were taken into consideration for the analyses. At the individual trials level, any values exceeding or falling below  $\pm 2.5$  standard deviations the individual mean for each experimental condition were treated as outliers and excluded (on average, excluded trial  $16\% \pm 2.16\%$  of the total). At the group level, participants exhibiting values that diverged by  $\pm 2.5$  standard deviations above or below from the group mean in each experimental condition were identified as outliers and omitted from the analyses (on average, excluded trial  $10\% \pm 3.26\%$  of the total). All significance tests were conducted with an  $\alpha$  significance level of 0.05. Post-hoc tests, when applicable, were carried out using the Bonferroni method.

To examine the relationship between performance in the joint action and visuo-tactile task, we conducted correlations between the RTs of the VTI and Grasping Asynchrony for the grasping task. To increase statistical power of this correlation, we pooled data from the three studies, including only the data from the common factors of Interaction Mode (Cued/Interactive), Movement Type (Same/Opposite) and Visuo-Spatial Congruency (Congruent/Incongruent). Specifically: 1) In the first study, only the trials in which the visual cue was presented on the front portion of the bottle (full yellow dot); 2) In the second study, only the data from the multisensory Visuo-Tactile condition; 3) In the third study, only the trials from the Late Timing condition were considered. We then created an index by subtracting the mean values of all Opposite Movement trials from Same Movement trials in Interactive Incongruent conditions. This index allowed us to assess whether the modulation of visuo-tactile interference effects elicited when coordinating with a partner to perform opposite compared to same movements was also associated to the ability to synchronously grasp the object during the joint action task. Statistical analyses were conducted using Statistica software (TIBCO Software Inc). Data, code, and materials are accessible upon request.

## **Experiment 1**

In our initial experiment, we explored whether the effects of multisensory integration during a grasping task were influenced not only by the participant's own movements but also by the actions

of a partner they are interacting with. Specifically, we examined whether when participant's needed to predict the virtual partner' actions in order to perform their own, a complementary interaction (where the partner's actions were directed towards different locations of the bottle shaped object) would enhance the individual's ability to detect tactile stimuli that are visually and spatially incongruent to his own action, but congruent to the one of the partner, compared to an imitative one (where the partner's actions were directed towards the same locations of the bottle shaped object), in which the visuo-tactile stimuli would result incongruent for both the participant and the partner.

### **Sample**

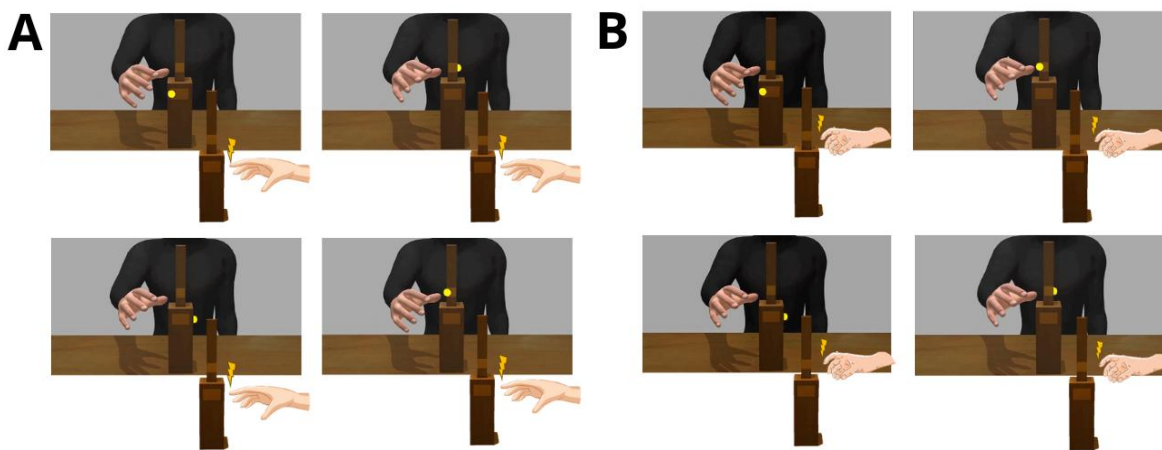
Reported effect sizes for differences between congruent and incongruent visuo-tactile multisensory effects in social and action tasks, range between  $\eta^2 = 0.12-0.35$  (Patanè et al., 2021; 2016; Senna et al., 2019). Accordingly, we conducted a power analysis using MorePower (Campbell, Thomson, 2012) assuming a  $\eta^2 = 0.21$ ,  $\alpha = 0.05$ , power = 0.8. This yielded to an estimated sample size of 32 participants. A total of 36 individuals (19 females; mean age 25.58 years, ranging from 18 to 36 years old; S.D. = 3.13) participated in the study, but 2 of them were identified as outliers in the VTI task, while 3 of them in the Joint Action Task and subsequently excluded from all analyses (see statistical analysis section above), resulting in a final sample of 34 participants for the VTI task and 33 for the Joint Action task.

A total of 32 participants were confirmed as right-handed based on the Standard Handedness Inventory (Briggs & Nebes, 1975), while the others were left handed. All participants had either normal or corrected-to-normal vision. Participants were initially kept unaware of the study's objective and were informed only after completing all experimental procedures. Written informed consent was obtained from all participant, in conformity to the ethical standards outlined in the Declaration of Helsinki. Participants were refunded 15 euros for their participation.

### **Design**

In this version of the human-avatar joint action task, experimental subjects were required to detect as quickly as possible where the tactile stimulation occurred (right hand thumb/index), while ignoring visual distractors. To report their answer, participants were instructed to use a keyboard, specifically: their left index finger to press the left arrow key (to indicate a tactile stimulus on their right index finger) or their left thumb to press the right arrow key (to indicate a tactile stimulus on their right thumb finger), depending on the location where they felt the stimulation on the other

hand. The visual cue was presented at the corresponding final position of the partner's right index finger (on the front of the avatar's bottle) or thumb (on the back of the avatar's object). The visual cue might appear slightly to the left of the virtual object, corresponding to the avatar's index finger on either the upper or lower portion of the bottle (Figure 2). Alternatively, a yellow half-dot could appear slightly to the right of the bottle, aligning with the avatar's thumb on either the upper or lower portion of the bottle (Figure 2). The distinction between a full dot and a half dot was intended to convey the impression that the dots were either appearing on the front (full dot) or on the back of the virtual bottle (half dot).



**Figure 2. Examples of Opposite (A) and Same (B) trials in Experiment 1.** There were not only congruencies and incongruencies between the visual cue's appearance along the bottle's vertical axis and the participant's grasping end-location (Visuo-Spatial Congruency Factor), but also congruencies and incongruencies between the stimulated finger (index/thumb) and the appearance of the yellow dot/half dot (front/back) (Visuo-Tactile Congruency Factor). Note that the pictures only show index stimulations, but also the thumb was alternatively stimulated in this first experiment.

The entire experiment comprised 480 trials, evenly distributed between Interactive (240 trials) and Cued (240 trials) interactive conditions, further divided between Opposite (120 trials) and Same (120 trials) movements. Within these, a subdivision was made between Visuo-Spatial Incongruent (60 trials) and Visuo-Spatial Congruent (60 trials) movements, reflecting the congruency between the participant's grasping end location and the location of the yellow dot/half-dot along the bottle's vertical axis (upper or lower portion). Moreover, there was an additional subdivision of these trials into Visuo-Tactile Congruent (30 trials) and Visuo-Tactile Incongruent (30 trials) stimulations, indicating the congruency between finger stimulation (index/thumb) and the dot's position along the bottle's depth axis (full/half dot). The 30 trials included Precision (10 trials) and Power (10 trials) grasping movements, and 10 different catch-trials: 6 trials with visuo-tactile cues delivered at

random time points within six equidistant intervals between 1/5th and 3/5th of the avatar's movement; 2 trials involving avatar course correction; 2 trials and with visual cues presented in the absence of tactile stimulation.

## Statistical Analysis

The dependent variables, analyses and inclusion criteria employed for the statistical analysis are the ones stated in the previous section (see Dependent Variables and General Statistical Analyses). In the following results sections only the significant effects of interests are extensively reported, while the other results are reported in tables.

### *Results of the VTI Task*

We conducted a within-subject Analysis of Variance (ANOVA) with the following factors: Interaction mode (Cued/Interactive), Movement type (Opposite/Same), Visual-Spatial congruency (VS-Congruent/VIS-Incongruent) and Visuo-Tactile congruency (VT-Congruent/VT-Incongruent) (Table 1).

**Table 1. Results of repeated Measures Analysis of Variance on RTs with Effect Sizes (partial eta-squared) reported.** Asterisks flags significant p-values.

Effect	SS	Degr. of	MS	F	p	Partial eta-squared	Non-centrality	Observed power (alpha=0,05)
Intercept	132891517	1	132891517	1406.02	0.00*	0.98	1406.02	1.00
Error	891517							
Error	3119037	33	94516					
Cued*Interactive	572440	1	572440	37.21	0.00*	0.53	37.21	1.00
Error	507641	33	15383					
Opposite*Same	8007	1	8007	9.10	0.00*	0.22	9.10	0.83
Error	29023	33	879					
VS Congruent*Incongruent	84669	1	84669	31.08	0.00*	0.49	31.08	1.00
Error	89892	33	2724					

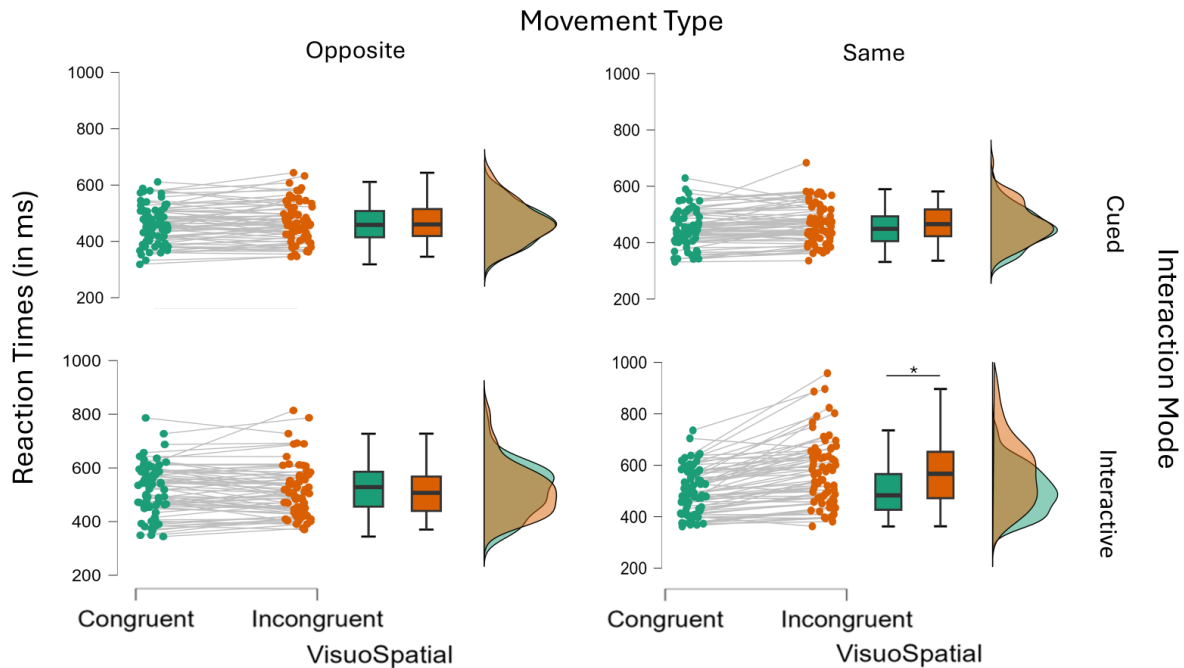
<b>VT Congruent*Incongruent</b>	<b>635 5</b>	<b>1</b>	<b>6355</b>	<b>4.71</b>	<b>0.04*</b>	<b>0.12</b>	<b>4.71</b>	<b>0.56</b>
Error	444 91	33	1348					
Interaction * Movement	199 35	1	19935	15.45	0.00*	0.32	15.45	0.97
Error	425 73	33	1290					
Interaction Congruency * VS	207 59	1	20759	18.72	0.00*	0.36	18.72	0.99
Error	366 04	33	1109					
Movement Congruency * VS	614 59	1	61459	33.41	0.00*	0.50	33.41	1.00
Error	607 09	33	1840					
Interaction Congruency * VT	96	1	96	0.14	0.71	0.00	0.14	0.07
Error	219 52	33	665					
Movement Congruency * VT	873 0	1	8730	13.19	0.00*	0.29	13.19	0.94
Error	218 40	33	662					
VT Congruency * VS	201 9	1	2019	2.22	0.15	0.06	2.22	0.30
Error	299 88	33	909					
<b>Interaction Movement Congruency * VS</b>	<b>370 62</b>	<b>1</b>	<b>37062</b>	<b>12.88</b>	<b>0.00*</b>	<b>0.28</b>	<b>12.88</b>	<b>0.94</b>
Error	949 65	33	2878					
Interaction * Movement * VT Congruency	114 7	1	1147	1.38	0.25	0.04	1.38	0.21
Error	274 81	33	833					
Interaction Congruency * VT Congruency * VS	515	1	515	0.84	0.37	0.02	0.84	0.14
Error	201 77	33	611					
Movement Congruency * VT Congruency * VS	2	1	2	0.00	0.96	0.00	0.00	0.05
Error	327 36	33	992					

Interaction * Movement * VT Congruency * VS Congruency	47	1	47	0.06	0.82	0.00	0.06	0.06
Error	278 94	33	845					

---

The 2x2x2 Interaction modes (Cued/Interactive), Visual-Spatial congruency (VS-Congruent/VIS-Incongruent), and Movement type (Opposite/Same) interaction was significant ( $F(1,33) = 12.88$ ,  $p < 0.01$ ,  $\eta^2 = 0.29$ ) (Figure 3). Post-hoc analyses revealed that the difference between VS-Congruent and VS-Incongruent stimuli was statistically significant only in the Interactive-Same condition ( $p < 0.01$ ), with no such significance in any other within-factor comparisons (all  $ps = 1.00$ ).

Specifically, it was observed that during the execution of Same movements in the Interactive condition, participants were faster to detect tactile stimuli during the Congruent Visuo-Spatial condition ( $M = 499.05$  ms;  $SD = 86.24$  ms) compared to the Visuo-Spatial Incongruent one ( $M = 574.12$  ms;  $SD = 131.94$  ms). This pattern is coherent with the notion that a visual cue appearing on a location different from the one of the individual's (and partner) action goal impairs the processing of a tactile stimulus. Crucially, this effect was not found during the execution of Opposite movements ( $p = 0.99$ ), suggesting that during interpersonal online interactions (Interactive mode), the interference on the discrimination of tactile stimuli generated by spatially Incongruent visual cues was selectively abolished by coordinating with a partner's movements that were spatially incongruent (Opposite) (Figure 3).



**Figure 3. Interaction Mode (Cued/Interactive) x Movement type (Same/Opposite) x Visuo-Spatial Congruency (VS-Congruent/VIS-Incongruent).** Asterisks indicate significant differences among conditions. The Tactile Response (in ms) was longer only in Same Incongruent trials, regardless of the Visuo-Tactile congruency between stimulated finger and visual cue position on the depth of the avatar’s bottle.

The Visuo-Tactile congruency factor (VT-Congruent/VT-Incongruent) exhibited a significant interaction with the Movement type factor (Same/Opposite) ( $F(1, 33) = 13.19, p < 0.01, \eta^2=0.29$ ) (Figure 4). Post-hoc analysis indicated that, during Same movement, Incongruent Visuo-Tactile trials resulted in a general interference effect ( $M = 505.51$  ms;  $SD = 81.98$  ms) compared to Congruent ones ( $M = 490.67$  ms;  $SD = 78.54$  ms;  $p < 0.01$ ).

Hence, the worsening of the performance (i.e., longer response times) was observed when the participant and the avatar executed the Same movement and the visuo-tactile stimulation was Incongruent at the finger level (i. e., if the participant felt the tactile stimulus on one finger but saw the visual cue close to a different finger of the partner). This result suggested that the dominant incongruity is that related to own finger and not the other's, that an individual is likely to be more affected by the incongruency related to their own finger rather than the partner's finger. This finding is in line with the study of Patanè et al. (2021), which also investigate the VT incongruency related to the individual’s fingers and not to the fingers of an observed agent executing the action.

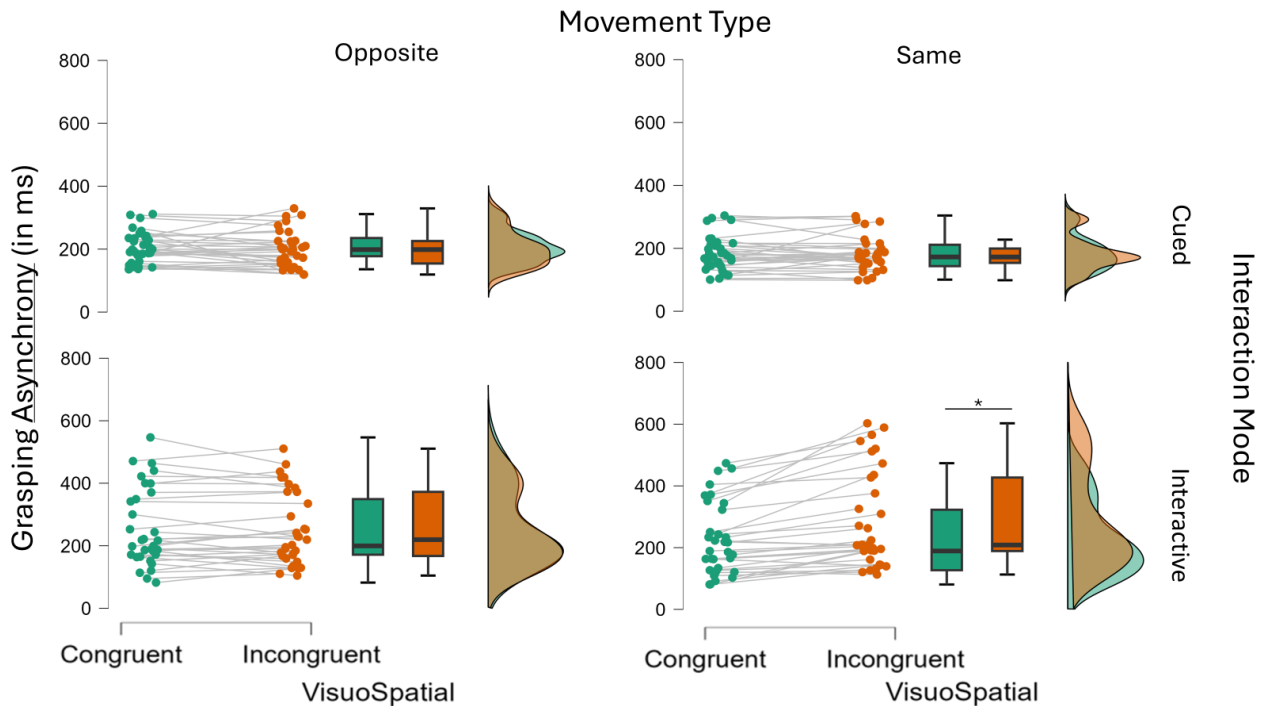


Interaction * VS Congruency	36748	1	36748	22.56	0.00*	0.41	22.56	1.00
Error	52118	32	1629					
Movement * VS Congruency	38236	1	38236	12.71	0.00*	0.28	12.71	0.93
Error	96277	32	3009					
Interaction * VT Congruency	2412	1	2412	1.76	0.19	0.05	1.76	0.25
Error	43921	32	1373					
Movement * VT Congruency	1002	1	1002	0.65	0.42	0.02	0.65	0.12
Error	49071	32	1533					
VT Congruency * VS Congruency	192	1	192	0.13	0.72	0.00	0.13	0.06
Error	48316	32	1510					
<b>Interaction * Movement * VS Congruency</b>	<b>22609</b>	<b>1</b>	<b>22609</b>	<b>10.64</b>	<b>0.00*</b>	<b>0.25</b>	<b>10.64</b>	<b>0.89</b>
Error	68017	32	2126					
Interaction * Movement * VT Congruency	1225	1	1225	0.94	0.34	0.03	0.94	0.16
Error	41522	32	1298					
Interaction * VT Congruency * VS Congruency	74	1	74	0.07	0.80	0.00	0.07	0.06
Error	34978	32	1093					
Movement * VT Congruency * VS Congruency	1747	1	1747	0.91	0.35	0.03	0.91	0.15
Error	61322	32	1916					
Interaction * Movement * VT Congruency * VS Congruency	200	1	200	0.18	0.67	0.01	0.18	0.07
Error	35335	32	1104					

Post-hoc analyses revealed that the difference between Congruent-Incongruent Visuo-Spatial stimuli was statistically significant only in the Interactive-Same condition ( $p < 0.01$ ), with no such significance in any other comparisons (all  $p$ s = 0.99) and in line with the findings from RTs in the VTI task.

Specifically, it was observed that in Incongruent Visuo-Spatial conditions participants were more asynchronous ( $M = 278.21$  ms;  $SD = 148.81$  s) compared to Congruent ones ( $M = 216.43$  ms;  $SD = 106.31$  ms) while performing Same movements in Interactive mode. This result was not found for the Opposite movements ( $p = 0.99$ ), suggesting that during interpersonal online interactions (Interactive mode), the interference generated by spatially Incongruent visual cues on the

discrimination of tactile stimuli had an effect also on the ability to coordinate with the partner (Figure 5).



**Figure 5. Interaction Mode (Cued/Interactive) x Movement type (Same/Opposite) x Visuo-Spatial Congruency (VS-Congruent/VIS-Incongruent).** Asterisks indicate significant differences among conditions. Grasping Asynchrony (in ms) is larger only in Same Incongruent trials, regardless of the Visuo-tactile Congruency between stimulated finger and visual cue position on the depth of the bottle.

It is also observed a main effect of the Visuo-tactile Congruency factor (VT-Congruent/VT-Incongruent)( $F(1,32) = 5.90$ ;  $p = 0.02$ ;  $\eta^2 = 0.16$ )(Table3). Specifically, it seems that for VT-Incongruent visuo-tactile stimulations, participants were significantly asynchronous ( $M = 217.70$  ms;  $SD = 102.37$  ms) than during the VT-Congruent ones ( $M = 210.95$  ms;  $SD = 97.48$  ms), in line with the findings from RTs in the VTI task.

### Accuracy

The  $2 \times 2 \times 2 \times 2$  Friedman ANOVA on Accuracy resulted significant ( $\chi^2 = 96.89$ ;  $p < 0.01$ ). We thus conducted Wilcoxon Matched Pairs Tests, correcting the significance level for (4) multiple comparisons ( $p < 0.01$ ) to mitigate the risk of false positives (Table 3). The analysis revealed significant differences in participants' accuracy depending on the congruency between visuo-spatial (VS) and visuo-tactile (VT) information across various movement conditions.

**Table 3. Wilcoxon Matched Pairs Test on Accuracy.** Here reported only significant tests ( $p < 0.01$ )

Variables	Valid	t	Z	p-value
Cued, Opposite, VS-Congruent, VT-Congruent *				
Interactive, Opposite, VS-Congruent, VT-Congruent	29	85.500	2.854	0.004
Cued, Same, VS-Incongruent, VT-Congruent *				
Interactive, Same, VS-Incongruent, VT-Congruent	27	56.500	3.183	0.001
Cued, Same, VS-Incongruent, VT-Incongruent *				
Interactive, Same, VS-Incongruent, VT-Incongruent	32	61.000	3.796	0.000
Interactive, Opposite, VS-Congruent, VT-Congruent *				
Interactive, Same, VS-Congruent, VT-Congruent	29	64.500	3.308	0.001
Interactive, Opposite, VS-Incongruent, VT-Congruent *				
Interactive, Same, VS-Incongruent, VT-Congruent	31	115.000	2.606	0.009
Interactive, Same, VS-Congruent, VT-Congruent *				
Interactive, Same, VS-Incongruent, VT-Congruent	27	29.000	3.844	0.000
Interactive, Same, VS-Congruent, VT-Incongruent *				
Interactive, Same, VS-Incongruent, VT-Incongruent	30	51.000	3.733	0.000

The analyses showed that participants performed significantly worse during Opposite movements when the visual cue appeared on the Congruent side relative to both the hand's final position on the bottle (VS-congruent) and the stimulated finger (VT-congruent), in the Interactive ( $M = 15.74$ ;  $SD = 2.08$ ) compared to the Cued condition ( $M = 16.94$ ;  $SD = 1.69$ ;  $p = 0.004$ ). In the Same movement condition, accuracy decreased significantly when the visual cue appeared on the incongruent side relative to the final hand position (VS-Incongruent) but congruent to the stimulated finger (VT-Congruent), in the Interactive ( $M = 14.71$ ;  $SD = 2.94$ ) compared to the Cued condition ( $M = 16.71$ ;  $SD = 2.43$ ;  $p = 0.001$ ). Performance further deteriorated when both in the visual-spatial and visual-tactile conditions the stimuli were incongruent ( $p < 0.01$ ) in the Interactive ( $M = 14.06$ ;  $SD = 2.85$ ) compared to the Cued condition ( $M = 16.56$ ;  $SD = 2.32$ ;  $p < 0.01$ ).

Within the Interactive mode, participants' accuracy was worse during Opposite movements ( $M = 15.74$ ;  $SD = 2.08$ ) compared to Same ( $M = 17.18$ ;  $SD = 2.21$ ) when both VS and VT were congruent ( $p = 0.001$ ). Interestingly and in line with the findings of the RTs, the performance reversely deteriorated when the visuo-tactile stimulation was congruent ( $p = 0.009$ ) in the execution of Same ( $M = 14.71$ ;  $SD = 2.94$ ) compared to Opposite movements ( $M = 15.97$ ;  $SD = 2.13$   $p = 0.009$ ). Accuracy also declined in Interactive-Same movements when: 1) the visuo-tactile stimulation was VT-Congruent for VS-Incongruent ( $M = 14.71$ ;  $SD = 2.94$ ) compared to VS-Congruent ( $M = 17.18$ ;  $SD = 2.21$ ;  $p < 0.01$ ); 2) when the visuo-tactile stimulation was VT-

Incongruent for VS-Incongruent ( $M = 14.06$ ;  $SD = 2.85$ ) compared to VS-Congruent ( $M = 15.79$ ;  $SD = 2.13$ ;  $p < 0.01$ ).

In summary, the results suggest that participants performed significantly worse when there were incongruencies between visual-spatial and visual-tactile information, especially during Same movements. The greatest accuracy decline occurred when both VS and VT cues were incongruent, though a notable reduction was also observed in conditions of VS congruence paired with VT incongruence.

## **Discussion**

This first experiment aimed at clarifying whether the effect of incongruent visual, spatial and tactile events VTI in one's own PPS and expanding to far visual stimuli during reaching and grasping movements (Brozzoli et al., 2009), is modulated by the movements of a partner when acting together. The results highlight that the spatial incongruency between participants' stimulated finger and the location of visual cue position on the partner's object led to a decline in performance (RTs and Accuracy) when detecting the tactile stimuli and in synchronising with the movements of the partner. Specifically, the interfering effect of the visual cue manifested only when it was necessary to predict the partner's actions (Interactive Mode), during imitative actions (Same Movement). In contrast, during complementary actions (Opposite Movement), the interfering effect of the visual cue was abolished when presented in an Incongruent space portion relative to where the subject needs to grasp, but congruent to the end location of the partner movement. This could occur because the participant was predicting the partner's action, whose end-movement position corresponded with the space portion where the incongruent visual cue was presented. These results seem to imply that during interpersonal interactions, visuo-tactile stimuli are influencing each other according to the partner's movement, suggesting that intersensory effects are due to an interpersonal motor interactive rule when performing joint actions.

When considering imitative movement in Congruent visuo-spatial conditions, the interferent VTI effect replicated and confirmed Patanè et al. (2021) findings. Indeed, if we attempt to "replicate" Patanè et al. (2021), conducting an ANOVA only on imitative (in which the avatar grasp the same portion of the object, as the participant, performing the same movement) Visual-Spatially Congruent trials (in which the visual stimuli would appear in the correspondent portion of the object grasped by the avatar and the participant), we replicated the established VTI effect during grasping actions (related to the correspondence between stimulated finger and the final finger

position on the object). The Incongruent Visuo-Tactile trials (with respect to the subject's stimulated finger)( $M= 511.60$ ;  $SD = 88.13$ ) were always interfered compared to Congruent ones ( $M= 486.49$ ;  $SD = 88.25$ ) ( $F(1,33) = 15.75$ ,  $p < 0.01$ ) (Figure 5). However, in our experiment it is important to notice that the visual cue was presented on the partner's object. This suggested that the VTI effects found in Patanè et al. (2021) for own objects were extended to the companion's one while interacting. However, in our experiment it is important to notice that the visual cue was presented on the partner's object. This suggested that the VTI effects found in Patanè et al. (2021) for own objects were extended to the companion's one while interacting. Hence, these results might shed light on the nature of low-level motor encoding of interactions in shaping intersensory effects and the social role of PPS.

## **Experiment 2**

To ascertain whether the observed visuo-spatial effects in Experiment 1 were not due to participants only paying attention to the visual cue and ignoring the tactile one, but rather reflected genuine intersensory interactions, we conducted a second study including a visual alone condition. Since vision is known to overshadow other sensory modality presented simultaneously (Rock & Victor, 1964; Spence & Driver, 2004), we intended to study whether asking to participants to detect the visual cue, or the tactile one would lead to similar pattern of results. To test this, we introduced a uni-sensorial Visual condition, in which a visual cue would be presented in absence of any tactile stimulation. Experiment 2 was aimed to investigate whether the incongruence effects for visual stimuli (unisensory Visual modality) would emerge alongside visuo-tactile stimuli (Visuo-Tactile modality).

### **Sample**

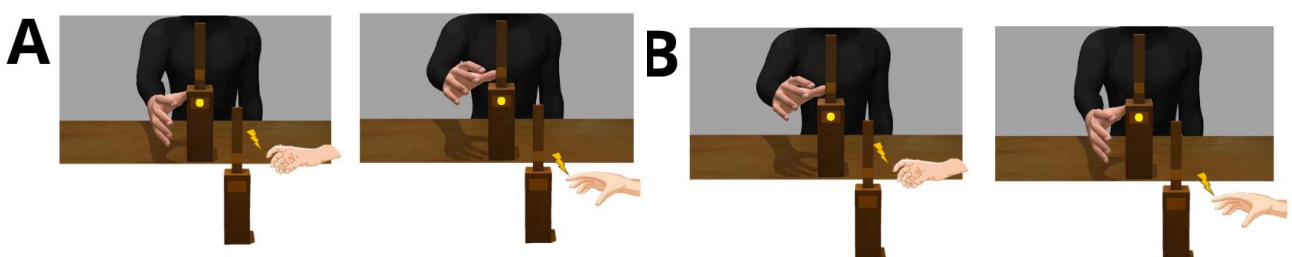
The choice of the sample size was selected through a statistical power analysis (Cohen, 2013) performed with the software More Power 6.0.4 (Campbell, Thomson, 2012). The expected effect size is the partial eta square ( $\eta^2 = 0.29$ ) value, from Experiment 1). The output indicates that in a  $2 \times 2 \times 2$  within factors design, a power of 0.8 and an eta squared of 0.29, requires a sample size of 22 participants. A total of 28 individuals participated in the study, but 4 of them were identified as outliers (see statistical analysis section above) and were subsequently excluded from all analyses. A final sample of 24 participants took part in the research. All participants (19 females; mean age 25.93 years  $\pm$ S.D. = 3.82) were initially unaware of the study's objective and were informed only

after completing all experimental procedures. Twenty-one of them were right-handed, based on the Standard Handedness Inventory (Briggs & Nebes, 1975), and all participants had either normal or corrected-to-normal vision. Written informed consent was obtained from all the individuals who took part in the study. The experimental protocols adhered to the ethical standards outlined in the Declaration of Helsinki.

## Design

The experiment consisted in a modified version of the task employed in Experiment 1, with two main conditions: (1) the Visuo-Tactile, in which the tactile stimulation, delivered on the dorsal part of participants' right index finger, was synchronised with a visual cue (a full yellow dot) appearing on the upper or lower part of avatar's bottle-shaped object; (2) the Visual, in which the visual cue would be presented in absence of any tactile stimulation (no white square would be presented on the bottom left of the screen to trigger the stimulation; see stimuli section). To respond to either the visual cue or the visuo-tactile stimulation, participants were required to press a foot pedal positioned and secured beneath the table with their right foot. Participants were required only to perform a detection task, since the thumb stimulation was not present and only the full yellow dot only appeared on the front of the avatar's bottle (in the upper/lower part) relative to Experiment 1 (Figure 6).

The experiment comprised a total of 192 trials. These trials were equally divided between Interactive (96 trials) and Cued (96 trials) conditions. Within these trials, half of them (48 each) required to perform the Same movement as the avatar, while the remaining half (48 each) the Opposite. Furthermore, half were Congruent (24) and half were Incongruent (24) trials within the Same and Opposite movement trials. Among the 24 Visuo-Spatial Congruent and Incongruent trials, 8 exhibited a power grip, 8 a precision grip, 2 presented modified trajectories with avatar direction changes, and 6 catch trials in which the visuo-tactile stimulation was delivered at six random intervals between the 1/5 and the 3/5 percentage of the avatar's movement start.



**Figure 6: Illustrations of the (A) Opposite and (B) Same movement types trials.** As depicted, the visual cue could result in a congruency or incongruency with the participant's grasping action, the virtual partner's action, or both (Visuo-Spatial Congruency).

## Statistical Analysis

The dependent variables, analyses and inclusion criteria employed for the statistical analysis are the same as in Experiment 1. In the following results sections only the significant effects of interests are extensively reported, while the other results are showed in tables.

### Results of the VTI Task

The following factors were considered: Interaction modes (Cued/Interactive), Sensory modality (Visual/Visuo-tactile), Visual-Spatial congruency (VS-Congruent/VS-Incongruent), Movement type (Opposite/Same) (Table 4). The analyses revealed a main effect of Sensory Modality ( $F(1,23) = 9.44$ ,  $p = 0.01$ ,  $\eta^2 = 0.29$ ), indicating significantly reduced RTs in the Visual condition ( $M = 324.32$ ;  $SD = 34.80$ ) compared to the Visuo-tactile one ( $M = 359.70$ ;  $SD = 22.23$ ).

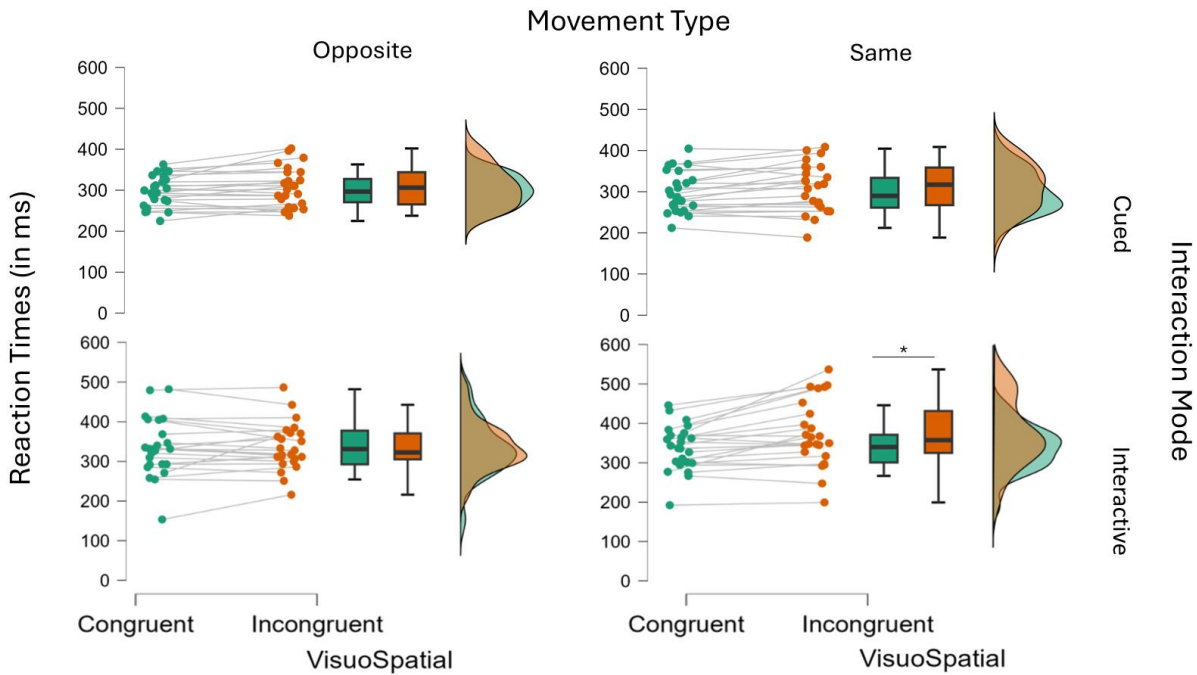
**Table 4. Results of repeated Measures Analysis of Variance on RTs with Effect Sizes (partial eta-squared) reported.** Asterisks flags significant p-values.

Effect	SS	Degr. of	MS	F	p	Partial eta-squared	Non-centrality	Observed power (alpha=0,05)
Intercept	40503505	1	40503505	905.52	0.00*	0.98	905.52	1.00
Error	1028775	23	44729					
<b>Visual * Visuo-tactile</b>	<b>82063</b>	<b>1</b>	<b>82063</b>	<b>9.44</b>	<b>0.01*</b>	<b>0.29</b>	<b>9.44</b>	<b>0.84</b>
Error	199988	23	8695					
Cued * Interactive	168269	1	168269	22.45	0.00*	0.49	22.45	0.99
Error	172384	23	7495					
Opposite * Same	13192	1	13192	10.18	0.00*	0.31	10.18	0.86
Error	29800	23	1296					
VS Congruency	21842	1	21842	23.97	0.00*	0.51	23.97	1.00
Error	20958	23	911					
Sensory Modality *	7765	1	7765	1.83	0.19	0.07	1.83	0.25
Interaction Error	97589	23	4243					
Sensory Modality *	2095	1	2095	1.37	0.25	0.06	1.37	0.20
Movement Error	35298	23	1535					
Interaction * Movement	7569	1	7569	7.19	0.01*	0.24	7.19	0.73
Error	24226	23	1053					
Sensory Modality * VS Congruency	8	1	8	0.01	0.93	0.00	0.01	0.05

Error			21153	23	920					
Interaction Congruency Error	*	VS	1355	1	1355	1.62	0.22	0.07	1.62	0.23
Movement Congruency Error	*	VS	7898	1	7898	7.17	0.01*	0.24	7.17	0.73
Sensory Interaction	Modality *		340	1	340	0.37	0.55	0.02	0.37	0.09
		Movement								
Error			20995	23	913					
Sensory Interaction Congruency	Modality *	VS	1690	1	1690	1.98	0.17	0.08	1.98	0.27
Error			19627	23	853					
Sensory Movement Congruency	Modality *	VS	1088	1	1088	1.13	0.30	0.05	1.13	0.17
Error			22219	23	966					
<b>Interaction * Movement * VS Congruency</b>			<b>6589</b>	<b>1</b>	<b>6589</b>	<b>5.87</b>	<b>0.02*</b>	<b>0.20</b>	<b>5.87</b>	<b>0.64</b>
Error			25816	23	1122					
Sensory Interaction VS Congruency	Modality *	Movement *	236	1	236	0.53	0.47	0.02	0.53	0.11
Error			10170	23	442					

The 2x2x2 interaction involving Interaction mode (Cued/Interactive), Movement type (Same/Opposite), and Visuo-Spatial Congruency (VS-Congruent/VIS-Incongruent), resulted significant ( $F(1,23) = 5.87$ ,  $p = 0.02$ ,  $\eta^2 = 0.20$ ) (Figure 8), replicating Experiment 1 findings. Post-hoc analyses revealed that the difference between Congruent and Incongruent Visuo-Spatial stimuli was statistically significant only in the Interactive-Same condition ( $p < 0.01$ ), with no such significance in any other within-factor comparisons (all  $p$ s = 0.99).

Specifically, it was observed that in Incongruent Visuo-Spatial conditions participants were slower ( $M = 374.11$  ms;  $SD = 85.26$ ) compared to Congruent ones ( $M = 337.91$ ;  $SD = 57.54$ ) while performing Same movements in Interactive mode (Figure 7). These results replicate the findings of Experiment 1. Importantly, the main effect of Sensory modality suggests that participants were actually following the instructions and showed faster RTs when detecting the visual cues compared to when detecting the tactile stimulus. However, the two modalities generated similar pattern of results suggesting that also tactile processing was affected by impending visual cues.



**Figure 7. Interaction mode (Cued/Interactive) x Movement type (Same/Opposite) x Visuo-Spatial Congruency (VS-Incongruent/VIS-Congruent).** Asterisks indicate significant differences among conditions. When participants needed to predict the partner’s movement, even if the stimuli was Visuo-Spatially Incongruent relative to their own end-position, they were as fast in detecting the stimuli as in Congruent conditions, regardless of the Sensory modality. This only happened for Opposite movements in Interactive modes, when the partner’s end-location was Congruent with the visual cue location. While, for Same movements, the classical VTI effect emerged when Incongruent visual/visuo-tactile cues were presented.

### Results of the Joint Action Task

The 2x2x2 interaction between Interaction modes (Cued/Interactive), Visual-spatial congruency (Congruent/Incongruent), and Movement type (Opposite/Same) yielded significant results ( $F(1,23) = 4.83$ ;  $p = 0.04$ ;  $\eta^2 = 0.17$ )(Table 5).

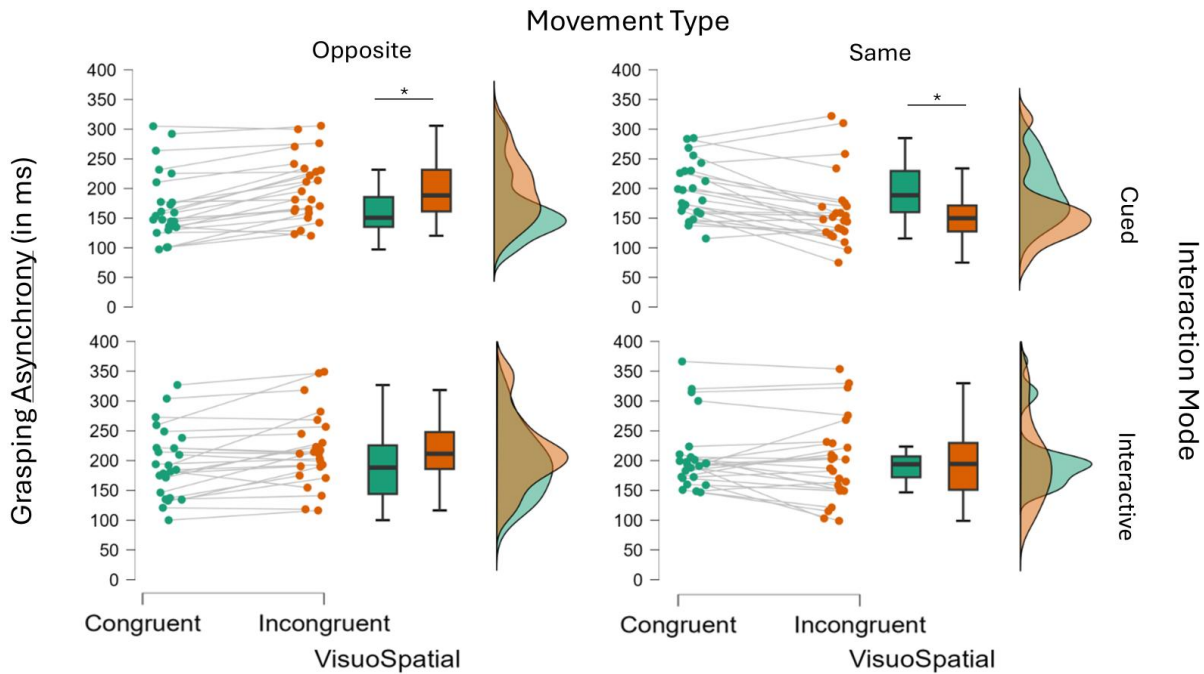
**Table 5. Results of repeated Measures Analysis of Variance on Grasping Asynchrony with Effect Sizes (partial eta-squared) reported (RTs).** Asterisks flags significant p-values.

Effect	SS	Degr. of	MS	F	p	Partial eta-squared	Non-centralit y	Observed power (alpha=0,05)
Intercept	14369919	1	14369919	381.86	0.00*	0.94	381.86	1.00
Error	865517	23	37631					
Visual * Visuo-tactile	2487	1	2487	0.42	0.52	0.02	0.42	0.10
Error	136392	23	5930					
Cued * Interactive	51675	1	51675	5.17	0.03*	0.18	5.17	0.59
Error	229913	23	9996					

Opposite * Same	1617	1	1617	0.72	0.40	0.03	0.72	0.13
Error	51480	23	2238					
VS Congruency	892	1	892	0.90	0.35	0.04	0.90	0.15
Error	22875	23	995					
Sensory Modality * Interaction	36	1	36	0.01	0.93	0.00	0.01	0.05
Error	102395	23	4452					
Sensory Modality * Movement	1998	1	1998	1.47	0.24	0.06	1.47	0.21
Error	31154	23	1355					
Interaction * Movement	66	1	66	0.04	0.84	0.00	0.04	0.05
Error	37134	23	1615					
Sensory Modality * VS Congruency	1363	1	1363	0.82	0.37	0.03	0.82	0.14
Error	38088	23	1656					
Interaction * VS Congruency	1943	1	1943	2.35	0.14	0.09	2.35	0.31
Error	19049	23	828					
Movement * VS Congruency	48924	1	48924	36.64	0.00*	0.61	36.64	1.00
Error	30710	23	1335					
Sensory Modality * Interaction * Movement	2565	1	2565	1.68	0.21	0.07	1.68	0.24
Error	35102	23	1526					
Sensory Modality * Interaction * VS Congruency	2089	1	2089	1.78	0.19	0.07	1.78	0.25
Error	26948	23	1172					
Sensory Modality * Movement * VS Congruency	16	1	16	0.02	0.88	0.00	0.02	0.05
Error	15130	23	658					
<b>Interaction * Movement * VS Congruency</b>	<b>6532</b>	<b>1</b>	<b>6532</b>	<b>4.83</b>	<b>0.04*</b>	<b>0.17</b>	<b>4.83</b>	<b>0.56</b>
Error	31106	23	1352					
Sensory Modality * Interaction * Movement * VS Congruency	600	1	600	0.83	0.37	0.03	0.83	0.14
Error	16720	23	727					

Post-hoc analyses, however, revealed that while no significant difference emerged between Congruent-Incongruent Visuo-Spatial stimuli in the Interactive condition ( $p > 0.05$ ), it is observed for both Opposite ( $p = 0.02$ ) and Same movements ( $p = 0.01$ ) in the Cued one.

Particularly, in Incongruent Visuo-Spatial conditions participants were more asynchronous ( $M = 199.00$  ms;  $SD = 54.21$ ) compared to Congruent ones ( $M = 169.63$ ;  $SD = 57.23$ ) while performing Opposite movements in Cued mode, suggesting that the classical incongruence effect was present when not predicting and online coordinating with the partner and had an effect on the ability to synchronise with the partner. In contrast, Incongruent Visuo Spatial trials ( $M = 163.24$  ms;  $SD = 61.11$ ) compared to Congruent ones ( $M = 195.52$ ;  $SD = 48.32$ ) resulted in significantly increased synchronization with the partner, suggesting that the incongruence effect when not predicting and online coordinating with the partner is mitigated, leading to improved synchronization despite the visuo-tactile spatial mismatch.



**Figure 8. Interaction Mode (Cued/Interactive) x Movement type (Same/Opposite) x Visuo-Spatial Congruency (VS-Incongruent/VIS-Congruent).** Asterisks indicate significant differences among conditions. Grasping Asynchrony (in ms) is larger in Opposite Incongruent trials only in Cued mode.

### Accuracy

For the Accuracy, no significant effect was found in the Friedman Anova ( $\chi^2 = 16.57$   $p = .35$ ).

### Discussion

The second experiment was designed to investigate whether the effects from Experiment 1 were intersensory in nature and to confirm the modulating role of the partner's movement on visuo-tactile integration processes. When the visuo-spatial motor condition was incongruent (e.g., the participant reaches for an object on the upper/lower portion and the visual cue appears on the opposite one) we would expect a reduced performance (RTs) in detecting a sensory stimulus, in both the unimodal and multimodal condition.

The results show that when there was incongruence between the participant's stimulated finger and the visual cue on the partner's object, performance in discriminating tactile stimuli declined (slower RTs). The visual cue's position on the partner's object influences the participant's tactile response only during imitative actions (Same Movement), modulating the VTI effects, but not

during complementary actions (Opposite Movement) when interacting. This was true for both the Visual and Visuo-Tactile sensory modalities, speaking in favour of the notion that the visual and tactile cues are integrated, otherwise a persistence of the visual cue interference on the ability to perceive a tactile stimulus during complementary movements was to be expected, rather than its integration.

Another result speaking in favour of a multisensory modulation in the Visuo-tactile condition, is the main effect of Sensory-modality, indicating a significantly reduced RTs in the Visual condition that could be interpreted in light of Fossataro et al. (2023) recent findings. In their findings, the authors report that when the visual cue is presented near the hand of somebody else in a visuo-tactile stimulation, the typical facilitation observed in multisensory conditions for congruent stimuli is abolished. This suggests that the proximity of another body of the visual cue can induce a 'shrinkage' of the PPS, thereby influencing the speed and efficiency of the sensory integration of asynchronously presented tactile stimuli. In our case, this framework could explain why the unisensory modality might enhance the efficiency of visual processing alone, compared to the multisensory one. Considering that our visual cues is presented on the other's object, this findings by Fossataro et al. (2023), might support our results, suggesting that the sensory modality's influence on RTs could be affected by the spatial positioning of stimuli relative to social co-agents action plan.

The general observed results pattern suggested that a participant's capacity to recognize visual cues during interpersonal interactions was influenced by the spatial congruency of these cues to both the participant's and the partner's movements. This implied that the predicting the partner's movements primed specific spatial locations, potentially heightening visual attention towards those areas. Importantly, increased visual attention was also manifested in the participant's tactile perception abilities. When continuously predicting the actions of others to perform complementary interactions (e.g., the agent reached for an object on one side and the partner reached for it on another side, where the visual cue appeared), congruency rules were modulated, so that an Incongruent visual stimulus with the respect to the action performed by the agent (e.g. the agent grasped the lower part of the bottle-shaped object and the visual cue appeared on the upper part of the partners' object) were processed as a Congruent stimulus, if the agent was predicting a complementary action of the partner (e.g. the partner was going to grasp the upper part of the bottle-shaped object, where the visual cue appeared). These findings seem to suggest that VTI is remapped by the rules of interpersonal motor interactions replicating the Experiment 1 results.

Furthermore, the observed results are not due to a visual dominance, but seem to be related to an actual integration of visual and tactile stimuli.

### **Experiment 3**

In the previous studies, participants received visual cues or visuo-tactile stimulations in a specific time window during the interaction. In this time window participants were able to predict the avatar's grasping end-location, as demonstrated in the pilot (see pilot section). In this experiment, we introduced another earlier stimulus presentation (where participants couldn't predict the avatar's directional goal) to the one implemented in the previous two experiments. Our hypothesis is that the facilitation of performance (RTs) in detecting tactile stimuli in incongruent trials during opposite movements—specifically, the abolishment of the visuo-spatial incongruency effect—would only occur in conditions where the avatar's reach-to-grasp end location is predictable.

#### **Sample**

The sample size was selected through the same statistical power analysis (Cohen, 2013) performed with the software More Power 6.0.4 (Campbell, Thomson, 2012) as in Experiment 2. A total of 20 individuals participated in the study, but 2 of them were identified as outliers, therefore subsequently excluded from all analyses (see statistical analysis section).

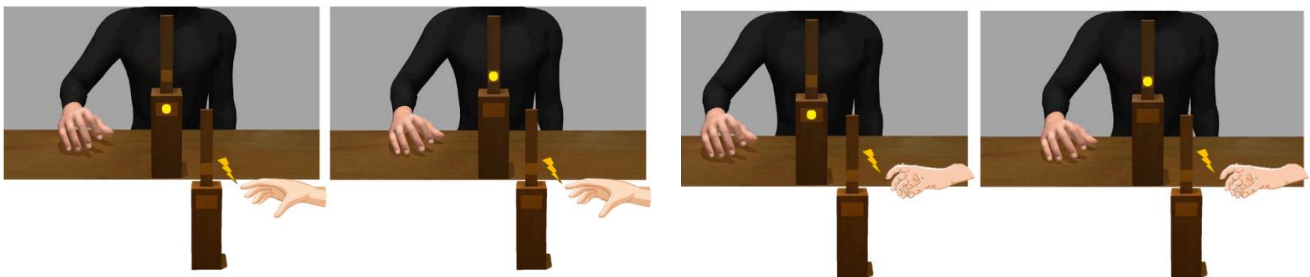
All participants (11 females; mean age 24.65 years, ranging from 21 to 30 years old; S.D. = 2.72) were confirmed as right-handed based on the Standard Handedness Inventory (Briggs & Nebes, 1975), and had either normal or corrected-to-normal vision. Participants were initially kept unaware of the study's objective and were informed only after completing all experimental procedures. Written informed consent was obtained from all participants in line to the ethical standards outlined in the Declaration of Helsinki.

#### **Design**

The procedure was identical to Experiment 1 and 2, except for the following changes. During the whole experiment participants would receive both the tactile and visual stimulation, except for the 10% of catch trials in which only the visual cue was presented and were not considered for the analyses. The tactile stimuli and response were administered as in the visuo-tactile sensory modality of Experiment 2. Based on the pilot results (see pilot section), the trials were equally

differentiated in two visuo-tactile stimulation timings: 1) Early presentation, before the aimed location of the avatar's movement could be detected, corresponding to the 1/5 of the avatar's movement start (Figure 9); and 2) Late presentation, when the avatars' movement direction could be accurately predicted, corresponding to the 3/5 of the avatar's movement start.

The experiment comprised a total of 336 trials. These trials were equally divided between Interactive (168 trials) and Cued (168 trials) conditions. Within these trials, half of them (84 each) required to perform the Same movement as the avatar, while the remaining half (84 each) the Opposite. Furthermore, half were Congruent (42) and half were Incongruent (42) trials within the Same and Opposite movement trials. Among these 42 trials, half presented an Early stimulation (21) and the other half a Late stimuli presentation. Within these 21 trials, 8 exhibited a power grip, 8 a precision grip, 2 presented modified trajectories with avatar direction changes, and 3 trials were catch trials without any tactile stimulation.



**Figure 9.** Example of different trials in which an Early visuo-tactile stimulation was delivered to the participants.

## Statistical Analysis

The dependent variables, analyses and inclusion criteria employed for the statistical analysis are the same as in Experiment 1 and 2. In the following results sections only the significant effects of interests are extensively reported, while the other results are showed in tables.

### *Results of the VTI Task*

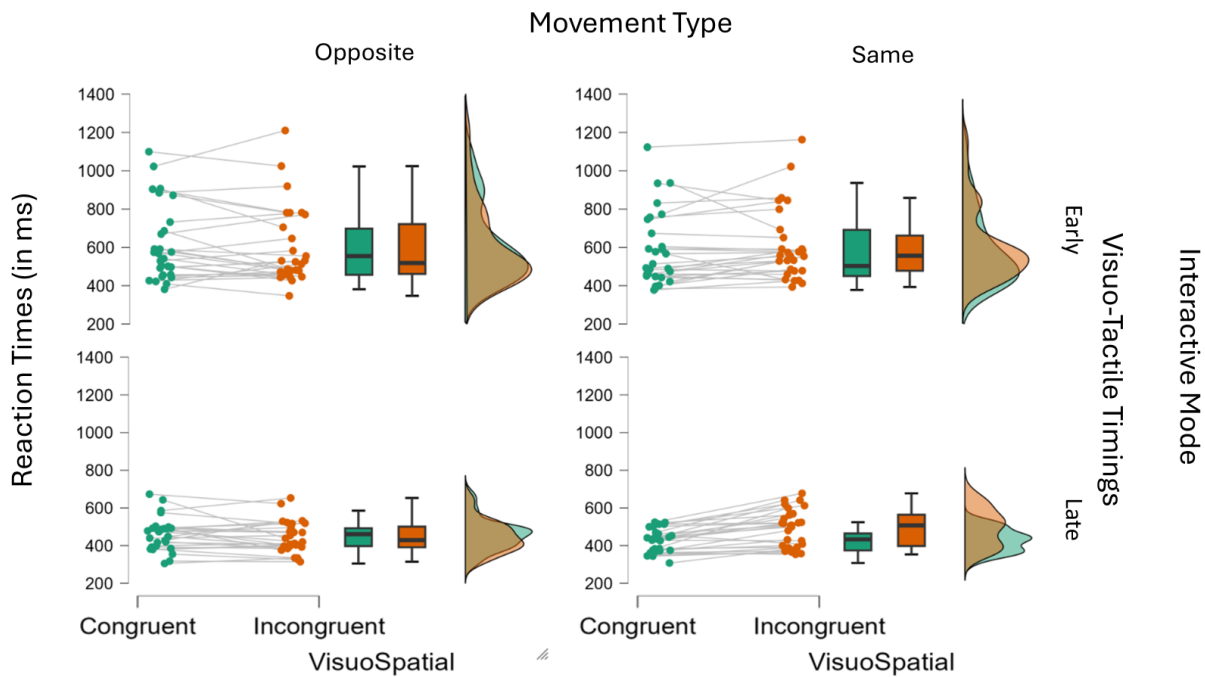
The structure of the within-subject 2x2x2x2 Analysis of Variance (ANOVA) on RTs comprehend the following Factors and Conditions: Interaction modes (Cued/Interactive), Movement type (Opposite/Same), Visual-Spatial Congruency (VS-Congruent/VS-Incongruent) and Visuo-Tactile Timing (Early/Late) (Table 6, Figure 10).

**Table 6. Results of repeated Measures Analysis of Variance on RTs with Effect Sizes (partial eta-squared) reported. Asterisks flags significant p-values.**

Effect	SS	Degr. of	MS	F	p	Partial eta-squared	Non-centrality	Observed power (alpha=0,05)
Intercept	58579968	1	58579968	1465.16	0.00*	0.99	1465.16	1.00
Error	679692	17	39982					
Cued*Interactive	188819	1	188819	30.21	0.00*	0.64	30.21	1.00
Error	106244	17	6250					
Opposite*Same	256	1	256	0.18	0.67	0.01	0.18	0.07
Error	23874	17	1404					
VS Congruent*Incongruent	1332	1	1332	0.55	0.47	0.03	0.55	0.11
Error	40983	17	2411					
<b>Early * Late</b>	<b>356186</b>	<b>1</b>	<b>356186</b>	<b>19.15</b>	<b>0.00*</b>	<b>0.53</b>	<b>19.15</b>	<b>0.98</b>
Error	316206	17	18600					
Interaction * Movement	697	1	697	0.36	0.56	0.02	0.36	0.09
Error	33074	17	1946					
Interaction * VS Congruency	51	1	51	0.06	0.82	0.00	0.06	0.06
Error	15779	17	928					
Movement * VS Congruency	35203	1	35203	9.46	0.01*	0.36	9.46	0.83
Error	63253	17	3721					
Interaction * VT Timing	5855	1	5855	0.92	0.35	0.05	0.92	0.15
Error	107697	17	6335					
Movement * VT Timng	480	1	480	0.27	0.61	0.02	0.27	0.08
Error	30573	17	1798					
VT Timing * VS Congruency	443	1	443	0.26	0.62	0.01	0.26	0.08
Error	29283	17	1723					
<b>Interaction * Movement * VS Congruency</b>	<b>8320</b>	<b>1</b>	<b>8320</b>	<b>7.71</b>	<b>0.01*</b>	<b>0.31</b>	<b>7.71</b>	<b>0.74</b>
Error	18335	17	1079					
Interaction * Movement * VT Timing	1257	1	1257	0.52	0.48	0.03	0.52	0.10
Error	40906	17	2406					
Interaction * VS Congruency * VT Timing	497	1	497	0.31	0.59	0.02	0.31	0.08
Error	27337	17	1608					
Movement * VS Congruency * VT Timing	287	1	287	0.12	0.73	0.01	0.12	0.06
Error	39827	17	2343					
Interaction * Movement * VS Congruency * VT Timing	268	1	268	0.19	0.67	0.01	0.19	0.07
Error	24121	17	1419					

However, the effects of Visual-spatial congruency (VS-Congruent/VS-Incongruent) and Movement type (Opposite/Same) only become significant once the partner's movement reaches a

point where the direction is clear, making the instructions Opposite/Same meaningful, and the Visuo-Spatial Congruency factor relevant. To ensure participants were executing the task correctly in the early visuo-tactile timing, i.e., responding promptly, we investigated the significant main effect of Visuo-Tactile Timing ( $F(1,17) = 19.15, p < 0.01$ ).



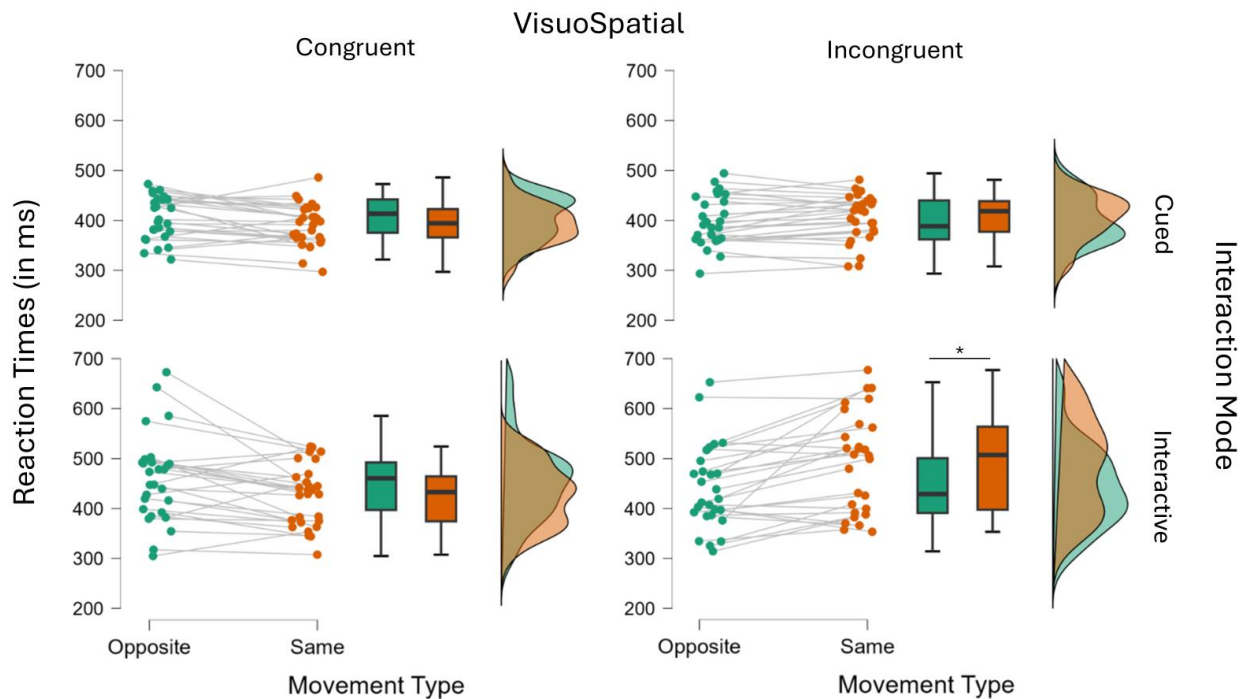
**Figure 10. Movement type (Same/Opposite) x Visuo-Spatial Congruency (VS-Congruent/VIS-Incongruent) x Visuo-Tactile Timings (Early/Late) in Interactive Mode . No significant difference was found in Tactile Response (in ms) between the Visuo-Tactile Timings.**

Reaction times were significantly slower in the Early timings ( $M = 486.17; SD = 69.46$ ) compared to the Late ones ( $M = 415.83; SD = 49.98$ ). These RTs indicated that participants waited before responding and their average answering time almost overlapped to the timing of the Late visuo-tactile stimulation ( $\approx 600$  ms from avatar movement start), instead of the Early one. Hence, our belief is that participants did not have fully adhered to the instructions, possibly due to the challenge of reacting to visuo-tactile cues at a very early stage of the reaching movement. The lack of interaction between the Visuo-Tactile Timing and the other factors should be interpreted cautiously, therefore the significant main effect of Visuo-Tactile Timing led to separate ANOVAs for Early and Late intervals, reported below (Table 7).

**Table 7. Results of repeated Measures Analysis of Variance on RTs with Effect Sizes (partial eta-squared) reported in Late Timings.** Asterisks flags significant p-values.

Effect	SS	Degr. of	MS	F	p	Partial eta-squared	Non-centrality	Observed power (alpha=0,05)
Intercept	24900	1	24900	1246.19	0.00*	0.99	1246.19	1.00
Error	33967	17	19981					
Cued*Interactive	64087	1	64087	20.71	0.00*	0.55	20.71	0.99
Error	52599	17	3094					
Opposite*Same	18	1	18	0.02	0.90	0.00	0.02	0.05
Error	17214	17	1013					
VS Congruent*Incongruent	119	1	119	0.15	0.70	0.01	0.15	0.07
Error	13423	17	790					
Interaction * Movement	1912	1	1912	1.50	0.24	0.08	1.50	0.21
Error	21633	17	1273					
Interaction * VS Congruency	434	1	434	0.42	0.53	0.02	0.42	0.09
Error	17564	17	1033					
Movement * VS Congruency	20926	1	20926	13.45	0.00*	0.44	13.45	0.93
Error	26452	17	1556					
<b>Interaction * VS Movement * VS Congruency</b>	<b>5789</b>	<b>1</b>	<b>5789</b>	<b>4.83</b>	<b>0.04*</b>	<b>0.22</b>	<b>4.83</b>	<b>0.55</b>
Error	20356	17	1197					

The 2x2x2 ANOVA on RTs (Interaction mode (Cued/Interactive), Movement type (Same/Opposite), Visuo-Spatial Congruency Relative to Own Action (Congruent/Incongruent) for Late visuo-tactile stimulation Timing (Figure 11), showed a significant interaction ( $F(1,17) = 4.83, p = 0.04, \eta^2 = 0.22$ ), thus replicating the findings of Experiment 1 and 2.



**Figure 11. Interaction mode (Cued/Interactive) x Movement type (Same/Opposite) x Visuo-Spatial Congruency (VS-Incongruent/VIS-Congruent) in Late Visuo-Tactile Timings.** Asterisks indicate significant differences among conditions. Results from experiment 1 and 2 were replicated: when participants were required to anticipate their partner's movements, they exhibited comparable response times in detecting the visuo-tactile stimuli, even when the stimuli are incongruent with their own anticipated final position on the bottle. This phenomenon was observed specifically in the case of Opposite movements, whereas for Same movements, the traditional VTI effect for incongruent trials was found.

Post hoc, however, revealed no significant difference between Congruent-Incongruent Visuo-Spatial stimuli in the Interactive condition for Same movements ( $p > 0.05$ ). Nevertheless, the Same Incongruent condition resulted significantly slower ( $M = 481.67$ ;  $SD = 100.61$ ) than the Opposite one ( $M = 431.61$ ;  $SD = 79.96$ ) in the Interactive mode, replicating the findings from Experiment 1 and 2, and confirming that VTI modulations are not found for complementary movements in interaction, where the incongruent avatar grip compensates for the incongruity of the visuo-tactile stimulation.

No significant interaction was found in the 2x2x2 ANOVA (Interaction mode (Cued/Interactive), Movement type (Same/Opposite), Visuo-Spatial Congruency Relative to Own Action (Congruent/Incongruent) for Early visuo-tactile Timing ( $p = 0.16$ ) (Table 8).

**Table 8. Results of repeated Measures Analysis of Variance on RTs with Effect Sizes (partial eta-squared) reported in Early Timings.** Asterisks flags significant p-values.

Effect	SS	Degr. of	MS	F	p	Partial eta-squared	Non-centrality	Observed power (alpha=0,05)
Intercept	34035937	1	34035937	881.73	0.00*	0.98	881.73	1.00
Error	656219	17	38601					
Cued*Interactive	130588	1	130588	13.76	0.00*	0.45	13.76	0.94
Error	161342	17	9491					
Opposite*Same	718	1	718	0.33	0.57	0.02	0.33	0.08
Error	37233	17	2190					
VS Congruent*Incongruent	1655	1	1655	0.49	0.49	0.03	0.49	0.10
Error	56842	17	3344					
Interaction * Movement	41	1	41	0.01	0.91	0.00	0.01	0.05
Error	52347	17	3079					
Interaction * VS Congruency	115	1	115	0.08	0.79	0.00	0.08	0.06
Error	25552	17	1503					
Movement * VS Congruency	14564	1	14564	3.23	0.09	0.16	3.23	0.40
Error	76629	17	4508					
Interaction * Movement * VS Congruency	2800	1	2800	2.15	0.16	0.11	2.15	0.28
Error	22100	17	1300					

### Results of the Joint Action Task

The 2x2x2 ANOVA conducted on Grasping Asynchrony (in ms) implying that the interplay between Interaction modes (Cued/Interactive), Movement type (Opposite/Same), Visual-spatial congruency (Congruent/Incongruent) separately for the Visuo-tactile Timing factor, showed no significance at the highest-order interaction for both Early ( $p = 0.49$ ) and Late ( $p = 0.55$ ) Timings (Table 9).

**Table 9. Results of repeated Measures Analysis of Variance on Grasping Asynchrony with Effect Sizes (partial eta-squared) reported . Asterisks flags significant p-values.**

Effect	Late Timings									Early Timings						
	SS	De gr. of	MS	F	p	Partial eta-squared	Non-centrality	Observed power (alpha=0,05)	SS	Degr. of	MS	F	p	Partial eta-squared	Non-centrality	Observed power (alpha=0,05)
Intercept	7778437	1	7778437	178.27	0.00*	0.91	178.27	1.00	6873245	1	6873245	159.06	0.00*	0.90	159.06	1.00
Error	741747	17	43632						734609	17	43212					
Cued*Interactive	37082	1	37082	2.86	0.11	0.14	2.86	0.36	194423	1	194423	7.72	0.01*	0.31	7.72	0.75
Error	22004	17	12944						42810	17	25183					

	6									9							
Opposite*Same	1664	1	1664	0.60	0.45	0.03	0.60	0.11		38	1	38	0.02	0.89	0.00	0.02	0.05
Error	46875	17	2757							34485	17	2029					
VS	6716	1	6716	3.14	0.09	0.16	3.14	0.39		6	1	6	0.00	0.95	0.00	0.00	0.05
Congruent*Incongruent																	
Error	36303	17	2135							24454	17	1438					
Interaction *	4060	1	4060	1.27	0.27	0.07	1.27	0.19		93	1	93	0.06	0.80	0.00	0.06	0.06
Movement Error	54152	17	3185							25231	17	1484					
Interaction * VS	14520	1	14520	9.02	0.01	0.35	9.02	0.81		1685	1	1685	0.72	0.41	0.04	0.72	0.13
Congruency Error	27356	17	1609		*					39773	17	2340					
Movement * VS	10386	1	10386	7.09	0.02	0.29	7.09	0.71		256	1	256	0.07	0.79	0.00	0.07	0.06
Congruency Error	24889	17	1464		*					58466	17	3439					
Interaction *	666	1	666	0.36	0.55	0.02	0.36	0.09		926	1	926	0.50	0.49	0.03	0.50	0.10
Movement * VS																	
Congruency Error	31120	17	1831							31316	17	1842					

### Accuracy

The Accuracy shows no significant effect in the Friedman Anova ( $\chi^2 = 10.73$ ;  $p = .77$ ).

### Discussion

The third study was aimed to replicate the findings of the previous ones and to confirm whether the representation of the PPS was modulated by the prediction of the movement that was about to be made while interacting. The present results showed that only for visuo-tactile stimulation presented at a timing in which participants were able to predict the partner's movement (Late), PPS seemed to be recruited. Specifically, VTI modulations are found only for the Interactive mode in Late visuo-tactile timings and not in Early ones. Again, the facilitatory effect for the detection of Congruent over Incongruent visuo-tactile stimuli was only present for imitative (Same) movements, compared to Opposite ones, in which was abolished.

Moreover, we observed that participants who took longer to detect tactile stimuli also showed delays in synchronizing their grasping actions with the avatar, indicating the impact of tactile processing speed on coordination.

These findings suggest that interactions and, specifically, the ability to anticipate the partner's movement, influenced the VTI effect here reported. PPS recruitment was involved only when the person was able to understand and predict the partner's movement direction, thus promoting the integration of one's own PPS representation with those of the partner. Furthermore, efficient tactile

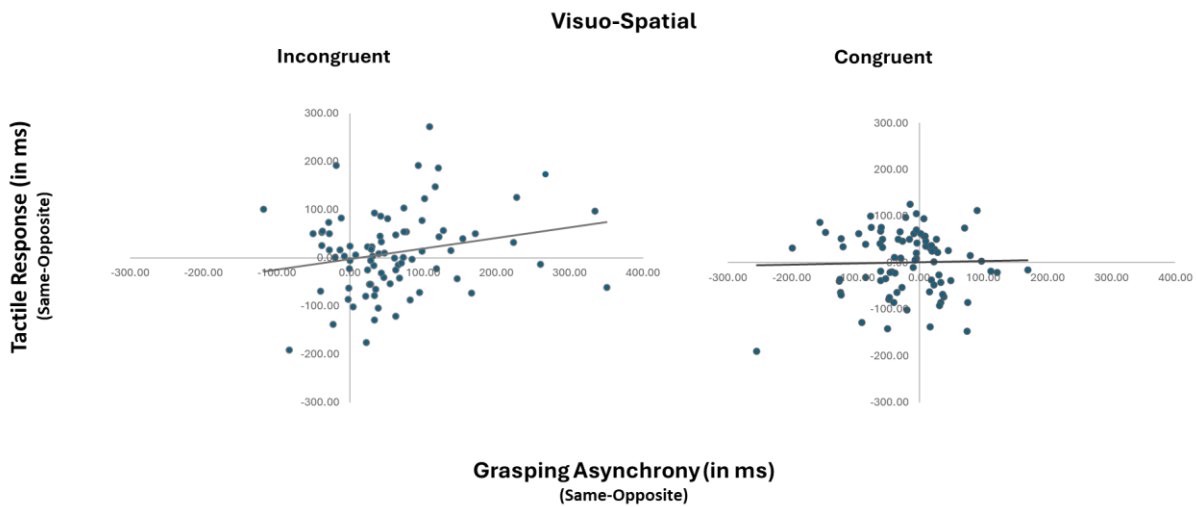
integration was found to be crucial for maintaining synchronized interactions during joint actions, especially when dealing with incongruent visuo-tactile stimuli.

These findings corroborate the previous results described in Experiment 1 and 2. Overall, these results supported our hypotheses that while actively interacting, multisensory mechanisms seemed to be related to a motor encoding and resonance of the partner's movement, allowing one to adapt and adjust to the other's action.

### **Relationship Between Visuo-tactile Interference Effects and Behavioral Performance at the Joint Action Task**

To examine the relationship between performance in the joint action and visuo-tactile task, we conducted correlations between the RTs (in ms) of Tactile Response and Grasping Asynchrony (in ms). To increase statistical power, we pooled data from the three studies, including only the data from the common factors of Interaction Mode (Cued/Interactive), Movement Type (Same/Opposite) and Visuo-Spatial Congruency (VS-Congruent/ VS-Incongruent) and created an index by subtracting the mean values of all Opposite Movement trials from Same Movement trials in Interactive mode separately for both VS-Congruent and VS-Incongruent conditions. This index allowed us to assess whether the modulation of visuo-tactile interference effects elicited when coordinating with a partner to perform opposite compared to same movements was also associated to the ability to synchronously grasp the object during the joint action task.

The correlation between participants' RTs to detect the tactile stimulation (Tactile Response) and in grasping the bottle together with the avatar (Grasping Asynchrony) in the overall three studies was found to be significant only in Interactive Incongruent conditions. Specifically, the analysis revealed that slower detection of tactile stimuli was associated with greater asynchrony in the joint grasping task ( $r(84) = 0.22$ ;  $p = 0.02$ ) when coordinating with a partner to perform opposite compared to same movements. This result is not found for the the Congruent Interactive condition ( $p = 0.81$ ) (Figure 12).



**Figure 12. Tactile Detection (Same - Opposite Movements) \* Grasping Asynchrony (Same - Opposite Movements) correlation across all three studies in Interactive (VS-Incongruent/VIS-Congruent) Visuo-Spatial condition.** Results from experiment 1, 2 and 3 were pooled and by subtracting the values (in ms) from Opposite movement trial to the Same ones we found a significant correlation between the two measures of Tactile Response and Grasping Asynchrony. Participants who exhibited significantly increased response times in detecting the visuo-tactile stimuli, also showed an enhanced asynchrony with the partner when the stimuli were Incongruent with their own anticipated final position on the bottle only in the Interactive task.

This indicates that participants who took longer to detect tactile stimuli also exhibited greater delays in synchronising their grasping actions with the avatar and viceversa. This finding, in line with previous discussed results, underscores the impact of VTI on coordination performance in the interactive tasks. When participants are slower in processing tactile information, it affects their ability to time their movements accurately with their partner. This related increase in the asynchrony of grasping suggests that efficient tactile integration is crucial for maintaining synchronised interactions. Furthermore, this is only true for Incongruent Visuo-Spatial condition. Thus, the ability to process and readily detect tactile stimuli is associated with the ability to coordinate with a partner during joint actions when processing Incongruent visuo-tactile stimuli and actively predicting the others' movements.

### 2.1.3. General Conclusion

These studies were designed to investigate whether the integration of visual and tactile information was influenced by the nature of interpersonal motor interactions. It is known that distant visual cues appearing on the action target location influence the ability to detect tactile stimuli that are delivered in spatially in/congruent locations within the PPS (di Pellegrino et al.,

1997; Spence et al., 1998; 2004). Crucially, research has shown that VTI incongruency effects, occurring during action execution, was also present even when the visual stimulus was distant from the body but aligned with the hand's endpoint position (Brozzoli et al, 2009; 2010; Patanè et al., 2018; Belardinelli et al., 2018; Senna et al., 2019). This effect persisted during action observation as well (Patanè et al., 2021). However, if and how these intersensory effects change during interpersonal interactions, where individuals need to online predict their partner's movements, remained unexplored.

Here we revealed that PPS recruitment emerged also when the visual cue appeared on the target object of an interacting partner and not one's own, only in conditions requiring continuous prediction and monitoring of the partners' movements. Interpersonal motor interactions may engage the neural processes associated with multisensory integration, related to bodily illusions (such as feeling sensations generated by the partner's movements within or on my body), and the integration of a tool (as the partner and its object). This integration suggested that the partner's movements and the object involved became instrumental in accomplishing the agent's goal, effectively serving as tools in his/her action. This would be translated into dynamic changes of the representations of one's own body and of the PPS, in order to facilitate sensorimotor predictions and consequently adapt one's own behaviour to the movements of the partner. Consistently with previous findings, our results contributes to support the notion that PPS recruitment is similar to a mirroring mechanism, wherein action-based somatosensory resonance is induced when observing others' actions (di Pellegrino et al., 1992; Brozzoli et al, 2013; Rizzolatti & Sinigaglia, 2016; Patanè et al., 2020). It would imply that the PPS recruits similar mechanisms in mapping one's own and the other's movements and represents a multisensory-motor interface that guides our interactions.

As predicted, our experimental results are in line with the previous literature on VTI incongruency effects (Brozzoli et al, 2009; 2010; Senna et al., 2019, Patanè et al., 2021), suggesting that sensory integration is influenced by the interpersonal interactions during joint actions. We highlight that the ability to predict the movement of one's partner is critical for the emergence of these modulations. Conversely, no multisensory modulation was found when participants did not need to predict the movement of the partner (i.e. our Cued mode). Although it is possible to speculate that this is due to different attentional effects and cognitive demands of the Interactive compared to the Cued condition on VTI, since the active prediction of the partner's movements necessitates continuous integration of one's own motor plan with the observed and predicted actions of the partner to achieve real-time coordination (Sebanz & Knoblich, 2009;

Knoblich & Jordan, 2003) and also introduces visuo-motor interference (Sacheli et al., 2012; 2013; Sacheli et al., 2015; Candidi et al., 2015; 2017; Curioni et al., 2017). Crucially, studies on cooperative joint-actions have shown that automatic imitation in the form of visuo-motor interference emerges when participants are performing complementary movements (Sacheli et al., 2012; 2013; Sacheli et al., 2015; Candidi et al., 2015; 2017; Curioni et al., 2017). Imitative interactions could be supported by a sensorimotor fronto-parietal simulative network (Kilner, Friston, & Frith, 2007; Urgesi et al., 2010; Avenanti et al., 2013), facilitating mutual adaptation and coordination. It is notable to consider that, in both the Cued and Interactive conditions, the execution of imitative (Same) and complementary (Opposite) movements were involved. Therefore, the Interactive condition might require greater cognitive and attentional resources, but mainly for the execution of complementary actions. If that was the case, only Opposite Interactive trials should have led to a higher demands of others' movements prediction, potentially inducing a stronger interference effect of the visual modality on the tactile one. However, at the behavioral level, it has been demonstrated that in this joint action task participants achieved an equal level of performance during complementary and imitative interactions (Era et al., 2018) in line with previous studies showing that complementary interactions are not more difficult than imitative ones (Ocampo & Kritikos, 2010; Sacheli et al., 2012; 2013). Looking specifically at our data across all three experiment we did not observe a main effect of Movement type (Same vs. Opposite) in the Grasping Asynchrony between the participant and the virtual partner grasping of the bottle-shaped object. This absence of significant differences suggests that the Interactive condition does not impose higher cognitive or attentional demands that would interfere with the coordination task performance in the complementary relative to imitative action. If anything, when considering the RTs related to the VTI effects, the incongruence effect primarily influenced the Same (e.g., imitative) movements in the Interactive condition. This finding is notable because, based on the aforementioned literature, imitative movements should generally facilitate coordination and reduce cognitive demand. Conversely, we did not find any worsening of RTs for the Opposite (e.g., complementary) movements, where a higher cognitive or attentional cost might be expected due to the above-discussed increased complexity of coordination due to action prediction and visuomotor interference mechanisms. These findings suggest that while the Interactive condition requires the continuous integration of predictive mechanisms to coordinate with the partner, VTI effects seemed to reflect the demands of the multisensory integration for a visual cue that results incongruent to both participants and the partner's end-movement position when detecting a tactile stimulus.

Moreover, the task implemented here allowed us to investigate how sensory events are integrated based on the interaction itself rather than individual action goals. These findings deepen our understanding of how multisensory information is integrated during joint action tasks and future studies may focus on exploring how interpersonal factors, personality traits and neural mechanisms impact VTI in social exchanges, fostering our understanding of low-level sensorimotor mechanisms in interpersonal tasks.

The results of the present study revealed an important impact of interpersonal coordination on multisensory motor processes, and opened up new questions about the value of sociality in PPS. This research contributed in clarifying how social interactions shape perception and action through the extension of personal space toward the other. These findings highlight the fundamental properties of social behaviour in shaping relatively low-level motor aspects of body-object interactions during the performance of (rather simple) actions and the profound role that sociality can have on individual perception.

## **2.2. Dyadic interactions based on multiple sensorimotor modalities modulate the behavioral and electrocortical markers of interpersonal cross-modal processing**

### **2.2.1. Introduction**

Sensorimotor Integration (SMI) is a process where sensory events are integrated with the motor commands that generated them. This integration may follow similar computational principles of Multisensory Integration (MSI, Jagini, 2021) and are crucial when integrating sensory information during a dynamic interaction with the environment. Indeed, the perception of sensory inputs is constantly probed and refined by movements performed by the agent via predictive mechanisms (Fuehrer et al., 2022). SMI is linked to the acquisition of so-called sensorimotor contingencies (SMCs), defined as learned relationships between movements and corresponding changes in sensory inputs (O'Regan and Noë, 2001). Although initially limited to perceptual processing at the individual level, SMCs are increasingly recognized as fundamental also in the understanding of processes occurring during effective interpersonal interactions. Specifically, recent works proposed to “socialize sensorimotor contingencies” (socSMC) and to investigate how the integration of sensorimotor signals from interacting partners might lead to the emergence of shared sensorimotor representations (Lubbert et al. 2021). In this framework, interpersonal motor interactions serve as a critical context where signals from multiple sensory channels are processed simultaneously, leading to cross-modal interpersonal integration.

At a pure sensory level, instead, Multisensory Integration (MSI) relies on the ability to weight sensory information from the same source according to optimal statistical rules (Ernst et al., 2002) to integrate them into the most reliable percept, as well as on the ability to estimate whether sensory information from different modalities originates from a common source and needs to be integrated or segregated (the so called “causal inference problem”; Kording et al., 2007). MSI and segregation are fundamental not only for representing external events such as those involving other people (Noppeney, 2021), but also for maintaining a stable and yet plastic representation of our own body, which is essential for efficient dynamic interactions with others (Blanke et al., 2015). Crucially,

sensory processing is affected during movement execution (via gating and attenuation, Kiltner & Ehrsson, 2022) and little is known about optimal MSI during movement execution.

To set the base for a sensorimotor grounding of interpersonal interactions, here we investigated whether engaging individuals into interpersonal interactions that exploited single (i.e., visual or tactile/proprioceptive) or combined (i.e., visuo-tactile/proprioceptive) sensory modalities would modulate the behavioural and electrocortical markers associated to cross-modal interpersonal processing.

### **2.2.2. Materials and Methods**

The set-up of the present project incorporates and adapts methodologies adopted by previous studies investigating visuo-tactile integration processes by means of EEG. In this section, I present information regarding the EEG setup and data, as well as in the next discussion section, to ensure completeness and clarity in the description of the study. It is important to note that while I have included these details, I did not personally conduct the analyses or the experimental portion associated with these data.

#### **Sample**

Twenty-five healthy subjects (15 males; mean age of  $24.8 \pm 2$ ) with any neurological or psychiatric condition reported were involved in the study. The sample size was determined through a prospective power analysis performed with the software More Power (Campbell and Thompson, 2012). We inserted as expected effect size the partial eta squared value (0.22) observed in Fossataro et al., 2023, where a similar cross-modal interpersonal set-up was deployed. The analysis indicated that a  $3 \times 2 \times 2$  within-subject design, a power of 0.90 and a partial eta squared value of 0.22, required a sample size of at least 24 participants. Participants' age varied between 20-29 years (mean age =  $24.8 \pm 2$ ). All participants were right-handed with normal or corrected-to-normal vision. Participants were naive as to the aim of the experiment and were informed of the purpose of the study only after all the experimental procedures were completed. The experimental protocol was approved by the ethics committee of the Fondazione Santa Lucia (Rome, Italy) and was carried out in accordance with the ethical standards of the 1964 Declaration of Helsinki and later amendments.

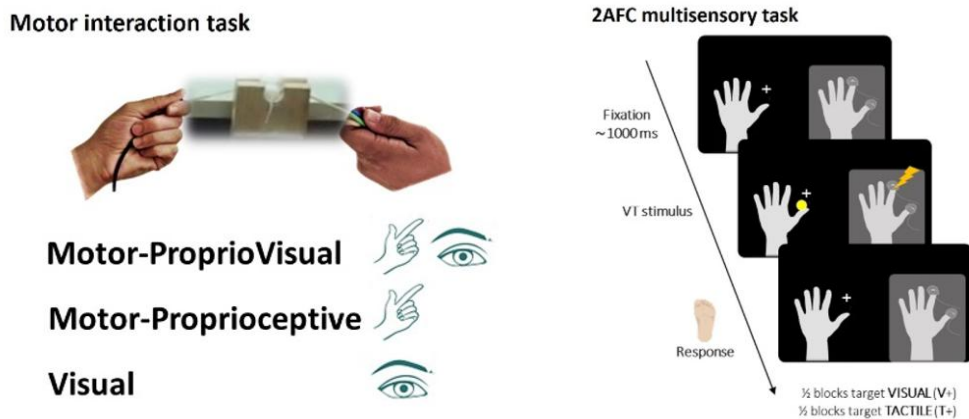
#### **Stimuli and experimental task**

Participants sat in front of a desk where a 1024 x 768 resolution LCD monitor laid horizontally at ~60 cm from their eyes. During the MSI task (see below), subjects were asked to place their left-hand palm down on the screen (11 cm to the left of their body midline), right (~15 cm) to the confederate's left hand, and to constantly fixate a white cross (1 cm diameter) positioned between the confederate's index and thumb (Fig. 1A). Tactile stimuli were delivered to the participant's left index/thumb using two 12 V round solenoids (0.8 cm in diameter) driving a magnet (1.5 mm in diameter) to the finger pad. Every time a current was delivered to the solenoid the magnet contacted the finger. The delivery of the current, and thus the presentation of the tactile stimuli, was controlled by a custom-built tactile controller (Heijo Research Electronics), which in turn was connected to the stimulus presentation computer. Although mechanical stimuli are less controllable than electrical ones in term of duration, each tactile stimulus was ~5 ms long, creating the sensation of a single tap. Before the beginning of the experimental procedure, subjects were explicitly asked if they could clearly feel the tactile stimuli. Visual stimuli consisted of flashes (1.2 cm in diameter) occurring next to the tip of the confederate's index/thumb. Visual and tactile stimuli were always presented simultaneously. Paradigm presentation and subjects' response were controlled and recorded using E-Prime 2.0 Professional software (Psychology Software Tools Inc., Pittsburgh, PA).

Specifically, they performed a candle-sawing activity (Soliman et al., 2015), that required the participant and the confederate each to use their left hand to bimanually and rhythmically operate a cutting wire sideways to saw a candle positioned horizontally on a custom-made holder. Importantly, the wire flexibility allowed participants to coordinate their movement in space (pulling the wire) and time (following and pacing the couple's movement rhythm) to reach the goal of sawing the candle (i.e., reach a joint goal). Moreover, such set-up allowed to isolate the proprioceptive channel (when subjects had to perform the task with their eyes closed – Motor-Tactile/Proprioceptive condition – M-P), the visual channel (when subjects had to passively observe the movement of the confederate – Visual condition - V), as well as to combine the two (when subjects had to synchronize their movement while observing the confederate's hand – Motor-ProprioVisual condition – M-PV).

In order to test the hypothesis that interacting with a partner by using one or two sensory modalities differently modulates the weight attributed to sensory information coming from one's own and the partner's body, after each interaction participants performed a cross-modal interpersonal, two-alternatives forced-choice, task (2AFC) (adapted from Maravita et al., 2002, 2003a). After the interaction, cross-modal interpersonal sensory integration processes were measured by means of

visuo-tactile interference effects realized on ones' own and the partner's hand (similar to Maravita et al., 2002, 2003a). Specifically, participants received tactile stimuli on their left hand (the same one they used during the interaction), either on the thumb or on the index, while the visual stimuli appeared close to the thumb or the index of the partner's left hand laying on a screen next to them (Fig. 1) and were asked to report as fast and as accurately as possible the location of the stimuli by pressing one of two foot pedals. White noise was played through headphones during the entire task, to attenuate EEG artifacts due to the noise produced by the mechanical tactile stimulators. Importantly, visual and tactile cues appeared in the same or opposite interpersonal somatotopic location (Congruency factor) with a 50% chance. Hence, in half of the trials the location of the visual and tactile stimuli differed (i.e. a flash appeared close the confederate's index while the participants received a tactile stimulus on her thumb – Congruency factor). Subjects were instructed to press a pedal either with their right or left foot, depending on the location of the target stimulus (left for stimuli on/close to the index and right for stimuli on/close to the thumb). Moreover, to test the role of top-down effects on multisensory integration processes, in half of the trials subjects would be asked to give their response based on the location of either the tactile stimulus (Tactile Relevance – T+) or the visual stimulus (Visual Relevance – V+ -Task Relevance factor). For each interaction (M-PVs, M-P, V), 4 consecutive experimental blocks were run. Each block began with an interactive session lasting 3 minutes, followed by a multisensory integration (MSI) measurement consisting of 40 trials, for a total of 160 trials per interaction (40 T+-Congruent, 40 T+-Incongruent, 40 V+-Congruent, 40 V+-Incongruent). At the beginning of the experiment, subjects performed two unisensory and one baseline session. During the formers, visual- (80) and tactile-only (80) stimulation was delivered, with subjects performing the same localization task as the MSI one described above. Importantly, visual and tactile stimuli were the same as the ones presented after each interaction. The latter would consist of 40 T+ and 40 T+ trials of the MSI task performed prior to any interaction, in order to measure cross-modal interference effects with no modulation related to the interactive session. The number of trials per condition was determined based on prior studies adopting similar paradigms targeting similar visuo-tactile neural processes (Sciortino et al., 2022).



**Figure 1. Experimental set-up.** Graphical representation of the interpersonal motor interaction task (left panel), and of the 2AFC visuo-tactile task (right panel).

### Electroencephalogram recording and processing

EEG signals were recorded and amplified using a gTec g.HIAMP Amplifier (g.Tec medical Engineering GmbH, Austria) and acquired from 126 active electrodes arranged according to the 10–10 system. Two electrodes were placed on the earlobes serving as reference, and all electrodes were physically referenced to the electrode placed on the right earlobe and then algebraically re-referenced off-line to the average of both earlobe electrodes. Impedance was kept below 30 K $\Omega$  for all electrodes for the whole duration of the experiment, amplifier hardware band-pass filter was 0.01-200 Hz and sampling rate was 1200 Hz. Offline, we applied a high pass filter to 0.1 and a low pass filter to 48 to remove slow drifts and line noise (50 Hz) from the signal. The two filters were applied separately as suggested by the latest version of EEGLAB (Delorme et al., 2004). Then, channels exhibiting excessive amounts of noise were detected by running the `clean_rawdata` function as implemented in EEGLAB (Delorme et al., 2004), by removing the channels which trace correlated with their direct neighbors for less than 80% of the time. After removing the above-mentioned channels, we epoched the continuous data into segments of 2400 milliseconds – from -1.200 to 1.200 around the onset of the stimulus. To identify and remove from the signal ocular artifacts (eye blinks and saccades) and excessive muscular noise, an Independent Component Analysis (ICA - Jung et al., 2000) was performed on the epoched data. Artefactual components were identified based on their time course and topography, and removed. On average, 9,9 (+2,75) components were removed per recording session (equivalent to 7,1% of the identified components per subject).

After the ICA, the bad electrodes previously removed (if present) were interpolated, with their signal being reconstructed based on the one from their surrounding neighbors with a “weighted”

approach, as implemented in Fieldtrip (Oostenveld et al., 2011). Finally, EEG data were re-referenced to the average of all channels, and epochs containing remaining artefacts such as electrode jumps were removed following visual inspection.

## **Data analysis**

### *Behavioural Data*

Participants' behavioral performance in the 2AFC task was quantified in terms of reaction times (RTs) and analyzed through 3 x 2 x 2, within-subject, repeated measures ANOVAs with Condition (PV/P/V), Task Relevance (T+/V+) and Congruency (Congruent/Incongruent) as within-subject factors (see extended methodology in the supplementary appendix). Only RTs from trials where subjects gave the correct response were taken into account for further analyses, with very few trials being inaccurate (average of inaccurate trials per subject = 2.4 + 1.6%). Also a repeated measures ANOVAs were performed, with Condition (M-PV/M-P/V), Task Relevance (T+/V+) and Congruency (Congruent/Incongruent) as within-subject factors. This allowed us to investigate whether the modulatory effect of a previous interaction on the participant's behavioural performance was general or, rather, specific either to the processing of multisensory incongruency or to the facilitatory effect elicited by congruent multisensory stimulation.

All statistical analyses were performed in Matlab r2022b and in the JASP Software. Post-hoc correction for multiple comparisons was conducted applying the Holm-Bonferroni method to the interactions that resulted as significant ( $p < .05$ ) from the ANOVA.

### *EEG Multivariate Analysis*

Single-subject EEG data were analysed through a multivariate pattern analysis (MVPA) to assess if and how the very same multisensory stimuli (i.e. multisensory processing) would elicit different neural processing patterns depending on the participant's previous interpersonal interaction. Multivariate analyses assume that the processing of different stimuli has different neural patterns associated to be exploited. Here, we investigated whether the very same multisensory stimuli would elicit different neural patterns depending on the observer's previous interaction. Analyses were performed using the MVPA-Light toolbox (Treder, 2020) in Matlab R2022b and custom-made scripts. Prior to every classification, data were normalized with a 200 ms prestimulus baseline, downsampled to 500 Hz and z-scored to center and scale the training data, providing numerical stability (Treder, 2020). Moreover, if a class had fewer trials than another, we corrected the imbalance by undersampling the over-represented condition (i.e., randomly removing trials). Lastly,

to increase the signal-to-noise ratio of the samples, training data (i.e. EEG trials) from the same class (i.e. PV, P, V) were split into groups of 5 and averaged.

First, we performed a multiclass decoding over time, investigating whether a Linear Discriminant Analysis (LDA) classifier was able to distinguish one condition (i.e. M-PV/M-P/V) from the other two based on the neural patterns in the time window from -200 to 1000 ms around stimulus onset. For linearly separable data, an LDA classifier divides the data space into  $n$  regions, depending on the number of classes, and finds the optimal separating boundary between them using a discriminant function to find whether the data fall on the decision boundary (i.e., 50% classification accuracy) or far from it (i.e., > 50% classification accuracy). A k-fold cross validation procedure in which samples were divided into 5 folds was repeated 20 times so that each sample was either used for training or testing at least four times, and the classifier's raw output was visualized to assess decoding performance over time. Statistical significance (i.e., performance significantly above chance level) was assessed through non-parametric cluster-based permutation tests (Maris & Oostenveld, 2007), using 'maxsum' for cluster corrections and Wilcoxon test for significance assessment. As a control analysis, we performed the same multiclass decoding, respectively, on T+ -, V+ -, congruent- and incongruent-only EEG trials. The classifier's raw output was visualized in order to assess decoding performance over time, and the decoding curves for each condition were compared through three dependent samples t-tests across time (M-PV vs M-P, M-PV vs V, M-P vs V). For the results of further control analyses please refer to the extended supplementary results. Both when extracting the time windows with performance significantly above chance level and the ones when the decoding curves differed between each other, statistical significance ( $p_{\text{clusterCorrected}} < .05$ ) was assessed through non-parametric cluster-based permutation tests.

Secondly, following a purely data-driven approach, we looked at the multiclass LDA's performance in a specific, early, time window in order to investigate whether the neural patterns differentiating the processing of multisensory stimuli after different interactions would be modulated already at low levels of the sensory processing cascade. To this aim, we performed a multiclass decoding, training and testing the classifier on the EEG data in the time window between stimulus onset (0 ms) and 150 ms. The accuracy values across testing folds of all repetitions were then averaged and presented on a confusion matrix to assess the probability of the classifier to in/correctly assign classes. As a control, we performed the same analysis on the EEG data in the time window between 150 and 300 ms post stimulus onset.

Lastly, we performed three binary classifications investigating how the performance of the classifier when distinguishing each condition from one of the other two evolved over time, as well as how

stable (generalizable) the neural representations contributing to such decoding were. Specifically, we adopted the temporal generalization method (King & Dehaene, 2014), training the classifier on each single time point and testing it at all time point in the same time window previously selected for the multiclass classification over time. This provided three temporal generalization matrices (time x time) with % accuracy values tested for statistical significance (i.e. performance significantly above 0.5 chance level) through non-parametric cluster-based permutation tests (Maris & Oostenveld, 2007).

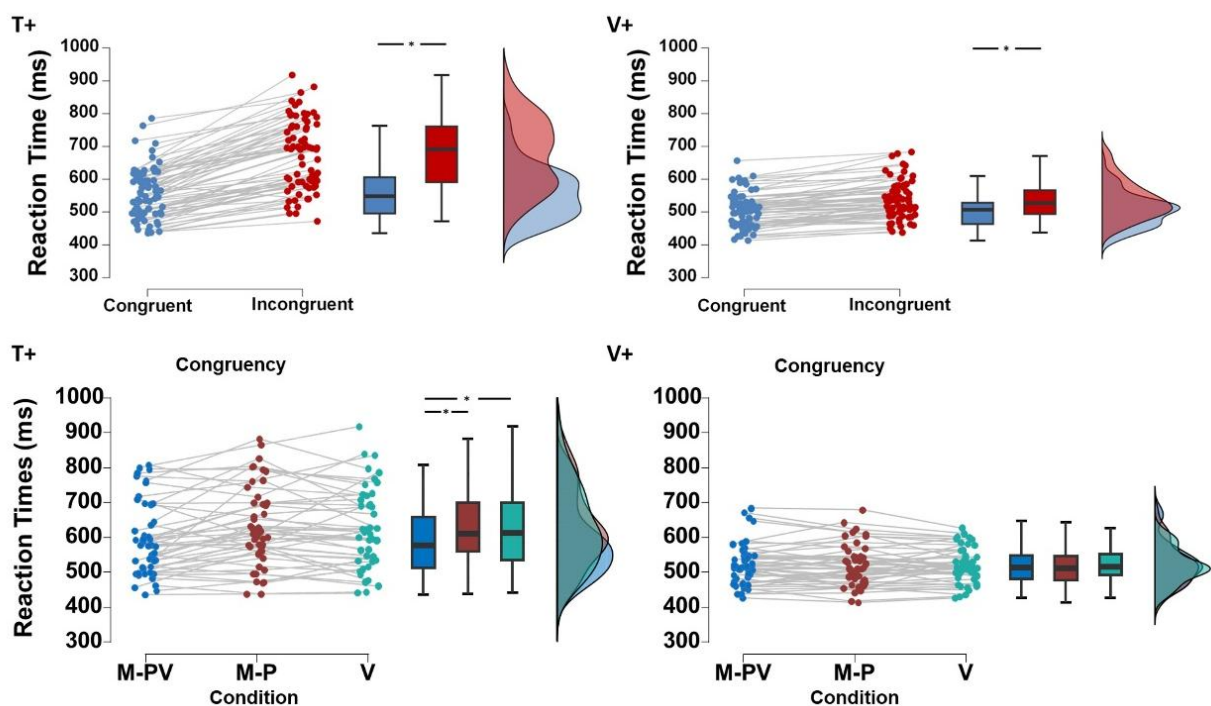
### **2.2.3. Results and Discussion**

Our results show that interpersonal motor interactions that rely on visual and/or motor-tactile/proprioceptive channels, modulate both the behavioral (i.e., Reaction Times) and neural (i.e., evoked EEG patterns) markers of cross-modal interpersonal integration.

First, our behavioral results extend the classical individual cross-modal interference effect (main effect of Congruency [ $F(1, 23) = 189.29, p < .001, \eta^2 = .89$ ]) of spatially incongruent stimuli to the case of interpersonal incongruency. Second, as already shown by previous studies on visuo-tactile perception (Goschl et al., 2014), subjects' performance was overall better when reporting the location of visual (i.e., V+ condition) compared to tactile (i.e., T+ condition) information (main effect of Task Relevance [ $F(1, 23) = 55.60, p < .001, \eta^2 = .71$ ]). Moreover, the interpersonal incongruency effect was stronger when reporting the tactile ( $p < .001$ ) compared to the visual ( $p < .001$ ) information (significant interaction between Congruency and Task Relevance [ $F(1, 23) = 71.24, p < .001, \eta^2 = .76$ ]), highlighting a smaller interferent effect of tactile over visual processing compared to the opposite (Figure 2, upper row plots).

Crucially for the present study, the performance of the subjects when focusing on touch (i.e., T+) was significantly modulated by the nature of the interpersonal interaction that preceded the cross-modal interpersonal integration task (significant interaction between Condition and Task Relevance [ $F(2, 46) = 5.8, p = .006, \eta^2 = .20$ ]). In details, subjects' reaction times significantly diminished

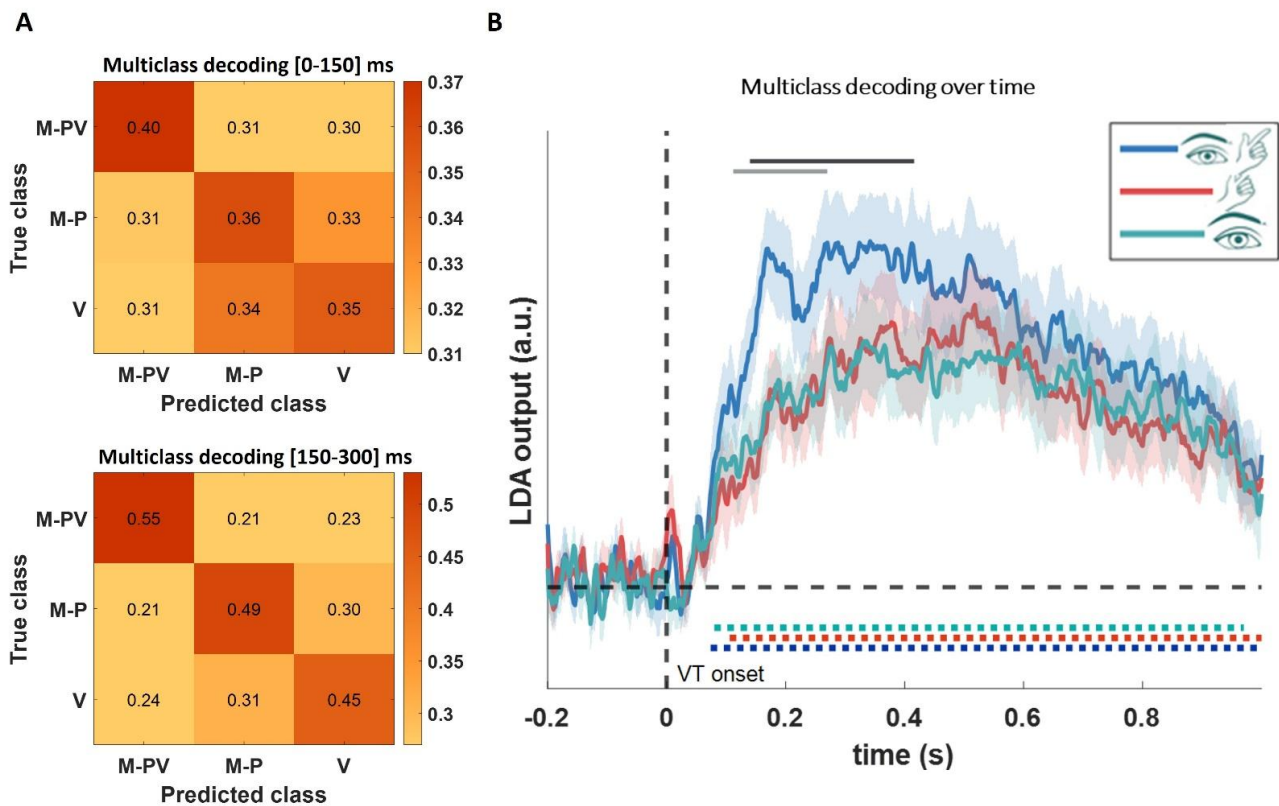
following the interaction involving both visual and tactile/proprioceptive modalities (i.e., the M-PV condition) compared to the two, single sensory modality, control conditions where either the visual (i.e., V;  $p = .017$ ) or the motor-tactile/proprioceptive (i.e., M-P;  $p = .002$ ) modality was used during the interaction (Fig. 2, lower row plots). This might be related to a modulation of the weights given to each sensory modality in order to reach optimal multisensory segregation and integration, respectively in incongruent and congruent trials (Ernst, 2002; Körding et al. 2007; Noppeney, 2021).



**Figure 2. Raincloud plots** showing the main effect of spatial congruency on the participants' reaction times (upper row), as well as the facilitatory role of the M-PV interaction on the performance of the participants when tactile location was task-relevant (lower row).

Second, our neurophysiological results complement this view, showing a clear effect of different interpersonal interactions on the neural representations of cross-modal interpersonal integration. Indeed, also the neural patterns evoked by identical visuo-tactile stimuli differ depending on the sensory channels involved in a previous interpersonal interaction. In detail, by adopting a multivariate approach, we show that a classifier is able to correctly identify the neural processing

associated with trials following a specific interaction as early as 74 ms after the onset of the stimulus (Fig. 3B), independently from the task relevance or the congruency of visual and tactile stimuli (see supplementary results). Moreover, such an early modulation of neural responses appeared to be particularly dependent on the more ecological interaction (i.e., the M-PV one) where participants had to integrate information coming from their motor-tactile/proprioceptive and visual channels in order to fulfill the joint task with a confederate. Indeed, when comparing the decoding curves for the M-PV condition compared to the M-P and the V one, the performance of the classifier resulted significantly better in a time window starting as early as 110 ms after stimulus onset (significant clusters: M-PV vs M-P = [110 270 ms],  $p_{\text{clusterCorrected}} = .005$ ; M-PV vs V = [140 416 ms],  $p_{\text{clusterCorrected}} = .007$  - grey lines on top of Fig. 2B). Moreover, as shown in the confusion matrices in Fig. 3A, when relying only on very early (0-150 ms) evoked patterns of activity, the classifier's capacity to distinguish the M-P from the V condition was worse than when it relied on later patterns (150-300 ms), while the M-PV trials were already clearly distinguishable in such an early time-window. This strengthens the behavioral results, suggesting that the interpersonal interaction where vision and motor-tactile/proprioception were to be functionally integrated, modulated differently the neural patterns related to the perception of following visuo-tactile stimuli.



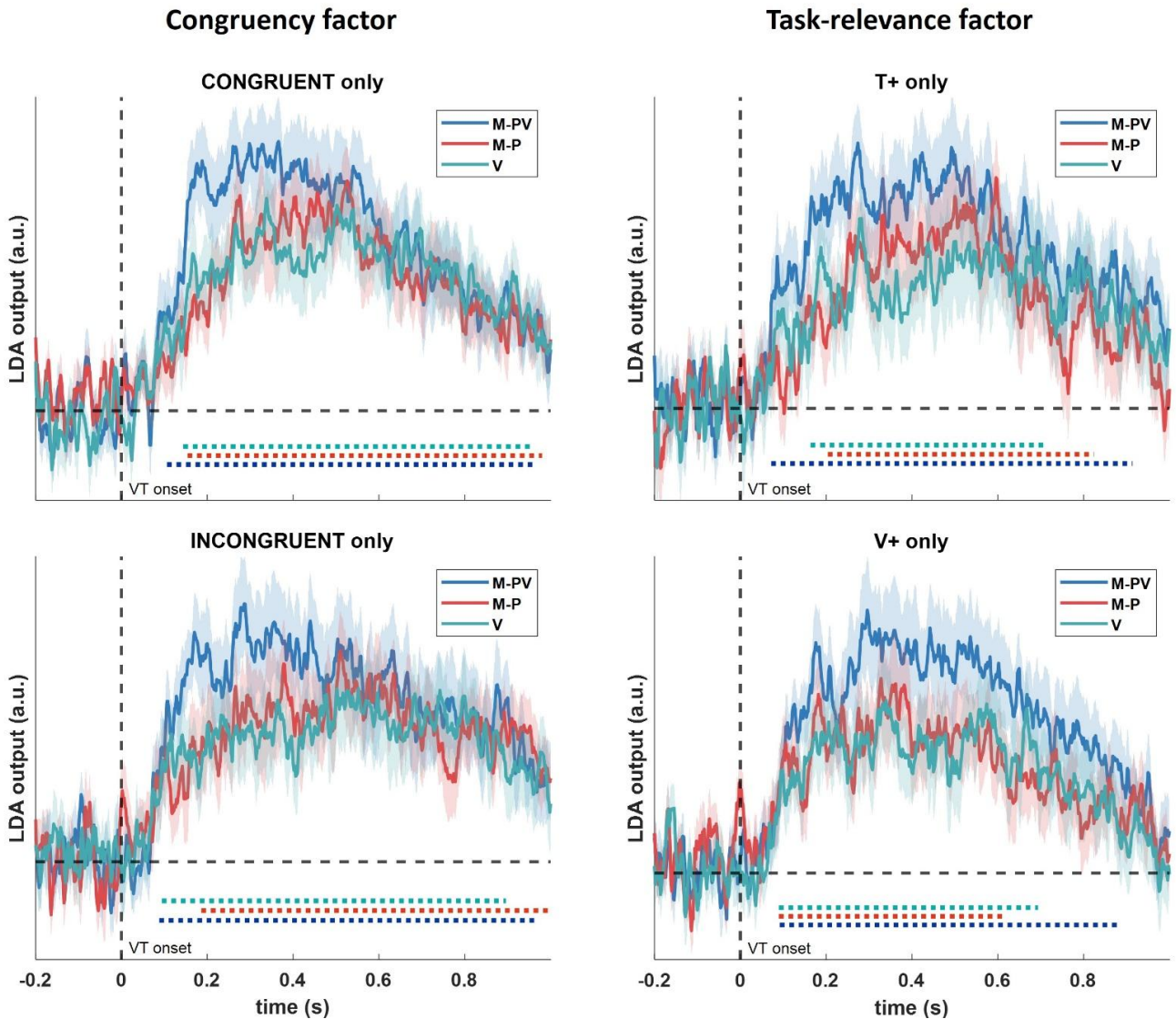
**Figure 3.** **A)** Confusion matrices showing the performance of the classifier in distinguishing the three behavioural interactions based on the neural patterns in the 0-150 ms time window (upper plot) and in the 150-300 ms time window (lower plot). **B)** Results from the multiclass decoding over time, showing that the classifier's performance was significantly ( $p_{\text{clusterCorrected}} < .05$ ) above chance level for all three classes (dotted lines at the bottom), as well as highlighting the significantly ( $p_{\text{clusterCorrected}} < .05$ ) better performance of the classifier in distinguishing the neural patterns following the M-PV interaction compared to the ones following the M-P one (light-grey line on top), and to the ones following the V one (dark-grey line on top).

To the best of our knowledge, these results are the first ones in the literature showing, 1) that cross-modal integration (Maravita et al., 2002) are present when sensory events happen close to ones' own and another person body, thus in an interpersonal (rather than individual) context, and 2) that using different sensory modalities to achieve interpersonal coordination induces an offline modulation of both behavioral and neurophysiological markers of cross-modal interpersonal integration. Furthermore, the present results extend previous literature on the influence of actions on sensory processes, where most of the studies focused on online effects occurring on a trial-by-trial basis (as the ones measuring adaptive responses to multisensory conflict - Noppeney, 2021). Most importantly, by taking advantage of MVPA, our work shows a temporally high-specific modulation of the distributed representations (across multiple regions) of visuo-tactile perception between one's own and a partner's hand following different interpersonal motor interactions. This could be

interpreted in light of previous results on the modulation of intersensory processing in the peripersonal space (PPS; Ronga et al., 2021) or by social factors (Teneggi et al., 2013). These first results require further investigations in order to be firmly interpreted in terms of an effect targeting specific computations (i.e., congruency processing, attentional selection). Nevertheless, they pave the way for new approaches to social neuroscience, focusing on plastic and time-lasting changes in behavioral responses to sensory stimulation and neural representations of intersensory effects depending on different contingencies between external sensory information (i.e., related to the environment and/or to other individuals) and internal motor and cognitive processing.

### **Supplementary Results**

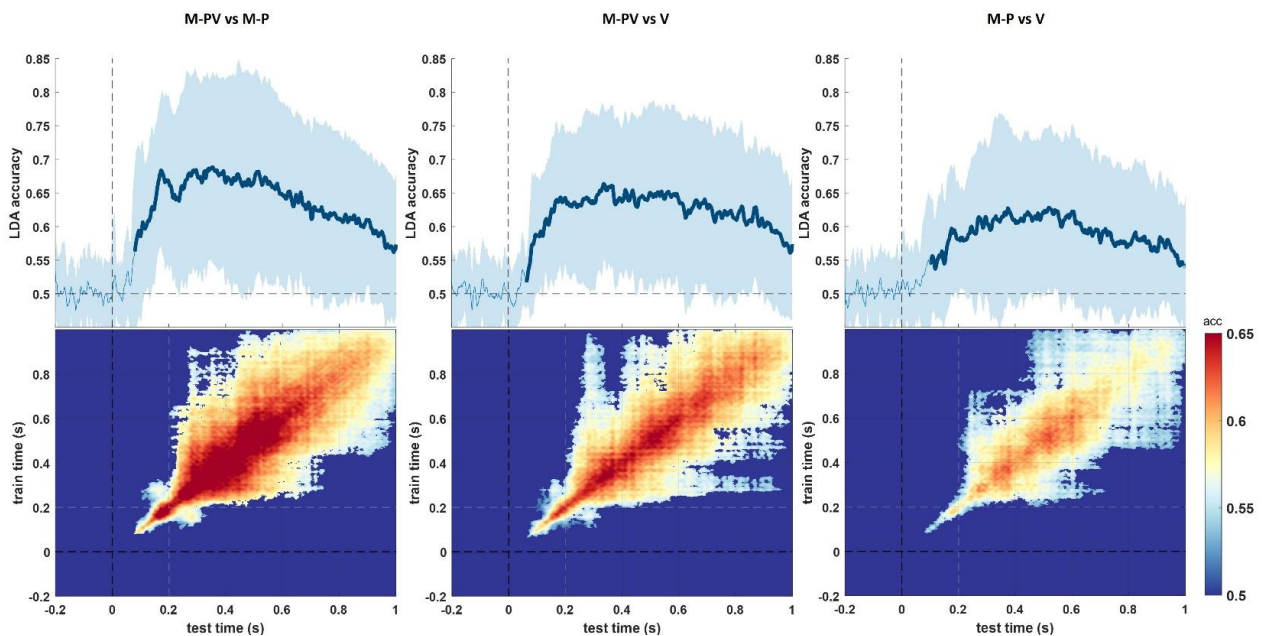
In order to check whether the classifier's performance in the multiclass decoding over time (Fig. 3B) was different if we did not collapse across the Congruency and Task-Relevance factors, we performed the same analysis separately for T+-, V+-, congruent- and incongruent-only data. The results are plot in Fig. S1, showing that, in all cases, the classifier's performance in distinguishing the three classes was significantly above chance level. Moreover, the M-PV condition was always better classified than the other two classes (which, in turn, lead to similar classification patterns).



**Figure S1.** Results from the control multiclass classifications over time, with the time points when the classifier’s performance was significantly above chance level highlighted at the bottom, for each class. Left column: the classifier decoded the Condition separately for congruent and incongruent EEG trials. Right column: the classifier decoded the Condition separately for T+ and V+ EEG trials.

Further, we performed three binary classifications investigating how the performance of the classifier when distinguishing each condition from one of the other two evolved over time, as well as how stable (generalizable) the neural representations contributing to such decoding were. Specifically, we ran three binary temporal generalization analyses. As proposed by (King & Dehaene, 2014), if a classifier generalizes from one-time window to another, then the underlying processes for the two-time windows would be similar, and the values in the resulting time x time plot would be represented as a single area spreading continuously out of the diagonal. If this is not the case, the results of the classification would be based on different information that travel across a

chain of transient neural representations, and the results in the time x time matrix would yield a diagonal generalization pattern. The results of this approach are reported in Figure S2. The time x time matrices, masked in order to show only significant results ( $p_{\text{clusterCorrected}} < .05$ ), highlighted a spread generalization pattern for each classification, suggesting that the neural patterns elicited by visuo-tactile stimuli are modulated by previous interactions in an interaction-specific and sustained-over-time way. Above each time generalization matrix (Fig. S2) we plotted the performance of the classifier along the diagonal, in terms of percentage accuracy. In detail, the classifier was able to discriminate visuo-tactile trials following the PV interaction from the ones after the P interaction from 80 ms post stimulus onset (significant cluster time window: [80 - 1000 ms]), PV trials from V trials from 64 ms post stimulus onset (significant clusters time window: [64 - 1000 ms]), and P trials from V ones from 100 ms post stimulus onset (significant clusters time window: [100 - 1000 ms]). These results are in line with the ones obtained from the multiclass decoding over time, highlighting a different modulation of early evoked activity by the M-PV interaction, compared to the other two, leading to earlier onsets of significant classification accuracy for the M-PV vs M-P and the M-PV vs V comparisons.



**Figure S2.** Results from the three binary classifications (M-PV vs M-P, M-PV vs V, M-P vs V). Bottom row: temporal generalization matrices, masked in order to show only significant accuracy scores ( $p_{\text{clusterCorrected}} < .05$ ). White dashed lines at 0.2 s were plotted in order to visually highlight the differences between the classifier's performances in the early time-window following stimulus presentation. Top row: classification performances over time (same as the diagonals in the time x time plots), with significant ( $p_{\text{clusterCorrected}} < .05$ ) accuracy clusters highlighted in bold.

## **2.3. Toward a “Social Rubber Hand”**

### **2.3.1. Introduction**

Body ownership and representation is a fundamental aspect of self-perception, also involved in social interaction and multisensory integration processes. The ability to distinguish between self and others relies on the integration of visual, tactile, and proprioceptive cues (Gallace & Spence, 2010; Ehrsson, 2020). This multisensory integration allows coherent body representation, essential for effective social interactions and bodily awareness (Blanke, 2012; Serino, 2019). Indeed, bodily awareness is the perception of one’s body in terms of position, movement, and ownership, achieved by integrating sensory input from vision, touch, and proprioception to distinguish the body from its surroundings (Bermudez et al., 1998).

The congruency of visuo-tactile and proprioceptive information seems determinant in the specific case of the rubber hand illusion (Walsh et al., 2011). Observing the touch on a rubber hand placed in a position compatible with one’s limb while simultaneously feeling a synchronous touch on one's hand can lead to a phenomenon known as the rubber hand illusion (Botvinick & Cohen, 1998). In this illusion, after a short period, individuals often experience the others’ hand as their own and show a proprioceptive drift of one’s hand toward the location of the rubber hand. Proprioceptive drift refers to the shift in perceived position of one's own limb towards the location of the rubber hand, and it is the most common measure used to evaluate the strength of the rubber hand illusion (Tsakiris & Haggard, 2005). This evidence supports the idea that bodily awareness relies on the integration of synchronous visuo-tactile information (Bermudez et al., 1998). It has been suggested that the rubber hand illusion could involve an interaction between body scheme representations and localized visuo-tactile integration (Tsakiris & Haggard, 2005).

Afferent tactile and efferent motor signals also contribute to self-recognition processes (Tsakiris et al., 2005). For example, when individuals engage in self-touch, tactile afferent perception tends to attenuate due to a ‘filtering’ effect of efferent signals associated with motor commands generated during self-initiated movements (Voss et al., 2006; Christensen et al., 2007). The possible contribution of being the author of an observed movement has been studied in inducing the embodiment of another hand; since agency appears crucial in building up the sense of ownership and achieving a coherent embodiment of the other hand in the rubber hand illusion (Tsakiris et al., 2006; Kammerts et al., 2009). Despite these insights, the role of interpersonal interaction in the

rubber hand illusion has yet to be explored. Therefore, in this study, we want to investigate the impact of interpersonal synchronous movement in eliciting embodiment and interpersonal intersensory effects in a novel paradigm that we would like to address as an attempt toward a «social» rubber hand illusion.

The process of anticipating the outcomes of a motor command is based on an efferent copy of it to forecast the consequences of the motor command and subsequently adjust the ongoing movement (Shadmehr & Krakauer, 2008; Miall & Wolpert, 1996). Interestingly, when engaged in an interaction, the internal motor model also predicts consequences for others' actions (Pezzullo, 2011; Friston, 2011; Wolpert et al., 2003). Indeed, an individual's efferent and afferent information both contribute to improving symmetrical interpersonal interactions (Pezzullo et al., 2017). Furthermore, simultaneous tasks may promote an internal representation that does not differentiate between the self and the other, aligning the actions and predictions of the agents (Pezzullo et al., 2017; Friston & Frith, 2015). To date, no studies have addressed the impact of social reciprocal interaction on the rubber hand illusion. When incorporating a mutual motor dynamic into this task, what effects does it yield?

It remains unclear whether this illusion extends to the hand of a partner during an interaction. To address this question, in the present experimental setup participants observed and executed stroking on a partner's hand while simultaneously being stroked on the exact location of their unseen hand by the partner. Participants underwent both synchronous and asynchronous (as control condition) interpersonal stroking blocks, in a counterbalanced order, where they interacted with a partner's hand positioned in front of them. Previous research indicates that the illusion fails to manifest if the rubber hand's position is unrealistic (Farnè et al., 2000; Tsakiris & Haggard, 2005; Kalckert & Ehrsson, 2012), as in this paradigm, in which it was rotated of 180° relative to the participants' hand. Particularly, the moving rubber hand illusion seems to rely on similar principles of temporal alignment and anatomical consistency as those found in the traditional rubber hand illusion (Ehrsson et al., 2004; Tsakiris & Haggard, 2005). It remains unclear whether proprioceptive drift would extend toward a partner's hand placed in an unrealistic position, as a result of the integrated input from both efferent and afferent signals during an interpersonal interaction.

Since Botvinik and Cohen (1998) demonstrated that the rubber hand illusion results from a complex interaction between vision, touch, and proprioception—where synchronizing visual and tactile stimuli alters the perceived position of one's hand—we aimed to explore this illusion's effects through a perceptual proprioceptive drift task (Durgin et al., 2007; Lloyd, 2007; Ehrsson et al., 2005; Farnè et al., 2000). After the stroking sessions, participants were asked to report the felt

location of their hand. We hypothesize that following synchronous stroking, participants will exhibit a greater proprioceptive drift towards the hand of the partner with whom they interacted. Additionally, we examined how the illusion affects interpersonal intersensory integration or segregation by employing a visuo-tactile crossmodal congruency task (Maravita et al., 2003). A tactile stimulus is presented on either participants' hidden left index or ring fingers. At the same time, a visual cue is presented near either the index or ring fingers of the partner's left hand. The hands of both the partner and participants, are respectively the ones stroked during the rubber hand illusion. We anticipated that synchronous mutual stroking would enhance visuo-tactile integration (VTI) effects, particularly for stimuli near the partner's hand, due to the shared sensory experience. Alternatively, we also predict that the asynchronous stroking sessions could lead to reduced VTI effects. This hypothesis is supported by the findings of Fossataro et al. (2023), which demonstrated that the proximity to others can induce a space representation shrink, effectively reducing the space in which multisensory responses occur. In the context of our paradigm, we suggest that asynchronous stroking, which lacks the temporal and speed congruence necessary for coherent multisensory integration, might exacerbate this boundary shift. Without synchronized tactile and visual stimuli, the brain may interpret the partner's hand as part of the 'other space,' rather than integrating it into a shared representation of space, thus leading to reduced VTI effects.

Moreover, since this process also involves top-down effects, participants' subjective report of rubber hand illusion (Botvinick & Cohen, 1998; Longo et al., 2008) and perceived self-other similarity (Myers & Hodges, 2011; Goldstein & Cialdini, 2007; Batson et al., 1997) are measured.

In summary, the present study aims to determine whether engaging in synchronous mutual stroking with a partner would alter interpersonal body schema representation and influence the dynamics of interpersonal intersensory integration/segregation. In addition, we aim to study whether this shared sensory experience affects higher-order social cognitive processes.

### **2.3.2. Material and Methods**

#### **Sample**

Reported effect sizes for differences between synchronous and asynchronous stimulation in the proprioceptive drift and questionnaire measures for the rubber hand illusion (RHI) range between  $\eta^2 = 0.19$ – $0.73$  (Gallagher et al., 2021; Aimola Davies et al., 2013; White et al., 2010; 2011). Based on the average of these effect sizes, we conducted an a priori power analysis using

MorePower 6.0.4 (Campbell & Thomson, 2012). The analysis indicated that a sample size of 28 participants would be sufficient to achieve a power of 0.85 and an  $\alpha = .05$ , with an effect size of 0.26. However, the final sample size was increased to 36 participants (mean age 22.08,  $\pm$  3.32) to ensure balanced conditions.

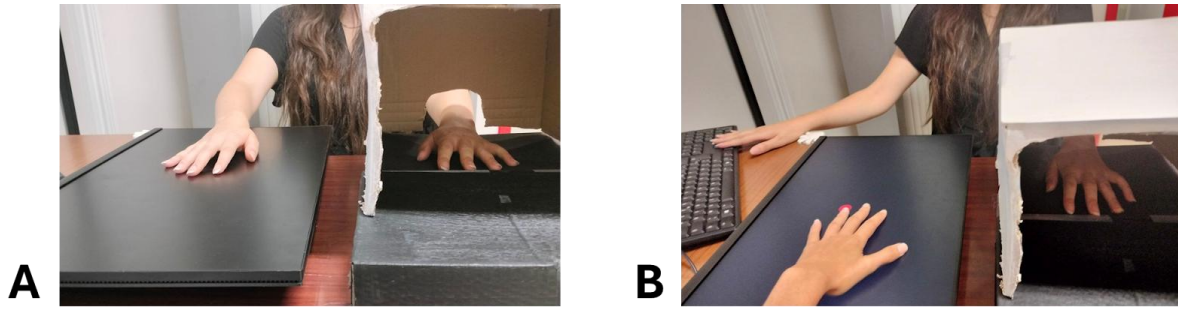
We tested only same-gender dyads, recruiting right-handed, healthy, French-mother-tongue females to match the experimenter's gender to minimize potential confounds related to gender-based differences in hand appearance.

Written informed consent was collected from participants before commencing the experiment. Initially, individuals were unaware of the study's objective and informed only after completing all experimental procedures. Participants were paid 10 euros per hour. The experiment was conducted in line with the Declaration of Helsinki, and approved by the local Ethics Committee of Université Paris Cité.

### **Set-Up**

Participants were tested in a quiet room. The experimenter, seated directly in front of the participants, acted as a confederate throughout the study. Participants' left hands were positioned inside a box measuring 20 cm in height, 36 cm in length, and 26 cm in width, accessed through a small hole (approximately 5 cm in diameter, allowing them to insert their arm. This hand position was fixed across participants and maintained throughout the experiment (hand relaxed, palm facing down, fingers pointing forward). The top of the box was covered to prevent visual feedback regarding hand position for the participants. In contrast, the side of the box closest to the experimenter remained open, enabling the experimenter to stroke the participant's hand during the 'social' rubber hand-stroking sessions (Fig. 1).

The experimenter's left-hand position mirrored that of the participants, approximately 30 cm to the left, and rotated 180°. The experimenter's left index fingertip was aligned on the same horizontal (x-axis) line as the participants' left index fingertip. Beneath the experimenter's left hand, a monitor was positioned measuring 5 cm in height, 54 cm in length, and 30 cm in width, setting the stage for the subsequent interpersonal intersensory task (IIT; described in details below). Under the participants' left hand, a little platform at the same height as the monitor balanced the hands' elevation (Fig. 1).

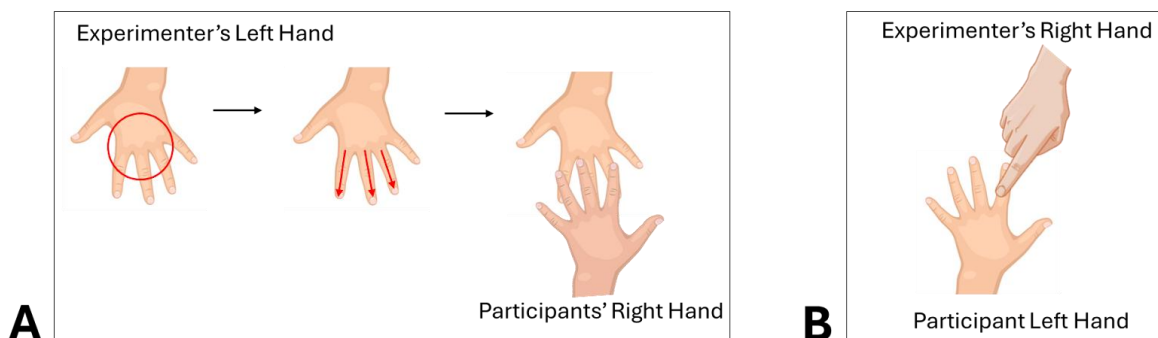


**Figure 1. Participants' hand positioning from the experimenter's point of view. (A)** Positioning before the stroking sessions of the 'social' rubber hand illusion. **(B)** Positioning for the IIT, also depicting where the experimenter's left hand was during both multisensory integration tasks.

### *'Social' Rubber Hand Illusion*

The strokings were performed with the experimenter and participants' right hands. Prior to the experiment, participants underwent a brief training session (approximately 1 minute) to learn the stroking technique. Participants were instructed to fixate on the experimenter's left hand, which was positioned directly in front of them (Fig. 1 - B) while executing each stroking using simultaneously the index, middle, and ring fingers of the right hand over the same fingers of the experimenter's left hand. The stroking movements were performed from knuckle to fingertip with a constant rhythm (Fig. 2 -A; around 1000ms).

During the Synchronous stroking sessions, the experimenter executed the same stroking movement as the participants regarding rhythm and touch location (Fig. 2 -A). In the Asynchronous stroking sessions, the experimenter stroked one of the three fingers (left index, middle, or ring) sequentially at a faster rhythm (Fig. 2-B). This approach was deliberately selected to avoid any inadvertent movement synchronization by the participants, which is crucial for preserving the integrity of the control condition and ensuring the validity of the results.



**Figure 2. Stroking Types.** (A) This panel illustrates the instruction and example of stroking performed by both participants and the experimenter during synchronous blocks. Participants were instructed to observe the experimenter's left hand and stroke it using simultaneously their right hand's index, middle, and ring fingers, caressing from knuckle to fingertip with a steady rhythm. (B) This panel depicts the stroking performed by the experimenter during asynchronous blocks on the participants' hidden left hand. Participants continued to stroke the experimenter's hand as in synchronous blocks. However, the experimenter stroked at a faster pace, sequentially touching the index, middle, and ring fingers of the participants' left hand, one at a time..

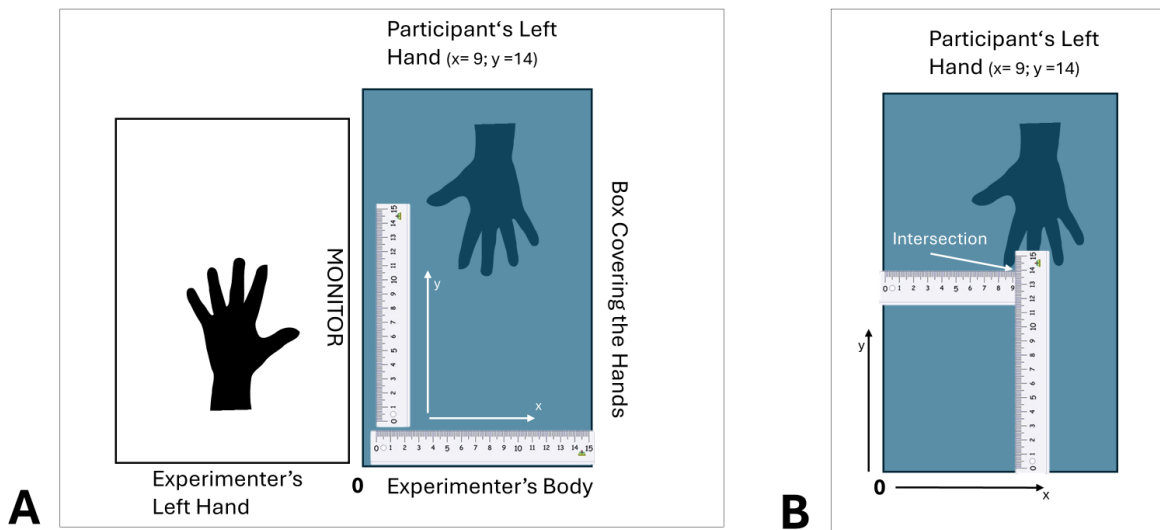
Both strokings were applied to the index, middle, and ring fingers to prevent potential fragmentation of body ownership (Tsakiris et al., 2006; 2005) and to mitigate specific influences during the subsequent IIT, which involved both the index and ring fingers. The order of synchronization for the stroking sessions was counterbalanced across participants, alternating between synchronous and asynchronous blocks. Each block included two stroking sessions, with each session lasting one and a half minutes.

## **Measures of the Illusion**

### *Proprioceptive Drift Task*

Before initiating the experiment (baseline) and after each stroking session (two per block), participants were asked to verbally report the perceived location of their hidden left index finger. In the literature, both the manual pointing task (Kalckert & Ehrsson, 2012; Riemer et al., 2013; Abdulkarim & Ehrsson, 2016; Abdulkarim et al., 2020) and the perceptual judgment task (Tsakiris et al., 2006; Kammers et al., 2009; Longo et al., 2008; Riemer et al., 2019) are commonly used to measure proprioceptive drift during the Rubber Hand Illusion (RHI). However, given that in our study both hands were engaged and influenced by the stroking procedure, we opted for a modified version of the perceptual judgment task to avoid any confounding effects from movement-related indications. Additionally, prior research has shown that perceptual judgments are more susceptible to the RHI than motor indication tasks (Kammers et al., 2009).

The experimenter, through two rulers moved along the horizontal and vertical axes above the box, would report the measurement. Participants vocally indicated their perceived index finger position at the intersection of the two rulers (Fig. 3), asking the experimenter to “stop” the rulers where they perceived the hidden hand to be. We considered both axes for measuring drift due to the interactive nature of our paradigm: the horizontal (X) axis represented drift towards the experimenter’s hand, while the vertical (Y) axis indicated drift towards the experimenter’s body.



**Figure 3. Representation of the proprioceptive drift task.** (A) Illustrate the set-up. The box's corner positioned on the midline of both the participant and the experimenter, on the side between the experimenter's body and hand, was considered the 'zero', the origin of the horizontal (x) and vertical (y) axes. The participants indicated vocally the felt location of the left hidden index fingertip with the aid of two rulers moved by the experimenter over the box along the horizontal (x) and vertical (y) axes from the origin. (B) Illustrate an example of an accurate judgment of the hidden finger. Participants were required to stop the rulers to indicate their index position at the intersection between the two rulers.

### *SRHI - Questionnaire*

Participants completed a questionnaire designed to assess the strength of the 'social' rubber hand illusion (SRHI-Q), which was adapted from a validated measure developed by Longo et al. (2008). This questionnaire evaluated the strength of the illusion within the experimental paradigm through three key components: Ownership, Location, and Hand-Loss. Although administered in French (Table 1 - A), the questionnaire is also presented here in English (Table 1 - B) for clarity.

Participants were instructed to complete the questionnaire after a stroking session once per block, reflecting on the stroking movements they had just performed with the experimenter. Consequently, participants filled out the questionnaire twice. It comprised 13 statements (listed below) measured on a 7-point Likert scale, ranging from "Strongly disagree" to "Strongly agree," with scores assigned from -3 to 3, respectively. Negative scores indicated disagreement with the statements, while positive scores indicated agreement.

The statements were administered through Qualtrics and presented in a random order. Statements 1 to 5 assessed Ownership over the experimenter's observed hand; statements 6 to 9 evaluated the perceived Location of the participants' hand and touch (both received and executed); and statements

10 to 13 gauged the sensation of Hand-Loss. The average scores for each set of statements provided scores for each component of the illusion (Table 1). We anticipated significantly increased scores in each component following synchronous blocks.

**Table 1. ‘Social’ RHI Questionnaire. (A) French Version, used in the experiment. (B) English back-translation.** Each group of statements referred to an illusion components that could occur during the stroking sessions: Ownership over the observed hand (1-5); Location of hand and touch (received and executed; 6-9); Hand loss sensation (10-13).

<b>A. French Version</b>	<b>B. English Translation</b>
Déclaration de propriété	
<b>1</b> Il m'a semblé que je regardais directement ma propre main gauche, plutôt que la main de ma partenaire.	<b>1</b> It seemed like I was looking directly at my own left hand, rather than the hand of my partner.
<b>2</b> Il m'a semblé que la main que je regardais commençait à ressembler à ma propre main.	<b>2</b> It seemed like the hand I was looking at began to resemble my real hand.
<b>3</b> Il m'a semblé que la main que je regardais m'appartenait.	<b>3</b> It seemed like the hand I was looking at belonged to me.
<b>4</b> Il m'a semblé que la main que je regardais était ma main gauche.	<b>4</b> It seemed like the hand I was looking at was my left hand
<b>5</b> Il m'a semblé que la main que je regardais faisait partie de mon corps.	<b>5</b> It seemed like the hand I was looking at was part of my body.
Déclaration de localisation	
<b>6</b> Il m'a semblé que je sentais le toucher exercé par ma partenaire à l'endroit où je voyais sa main gauche.	<b>6</b> It seemed like I was feeling the touch made by my partner in the location where the hand I was looking at was.
<b>7</b> Il m'a semblé que ma main gauche semblait être là où était la main que je regardais.	<b>7</b> It seemed like my left hand was in the location where the hand I was looking at was.
<b>8</b> Il m'a semblé que la main que je regardais était là où se trouvait ma main gauche.	<b>8</b> It seemed like the hand I was looking at was in the location where my left hand was.
<b>9</b> Il m'a semblé que le toucher que je ressentais était causé par ma main qui touchait la main de ma partenaire.	<b>9</b> It seemed like the touch I felt was caused by me touching the hand I was looking at .
Déclaration de la perte de main	
<b>10</b> Il m'a semblé que je ne pouvais pas bouger ma main gauche.	<b>10</b> It seemed like I was unable to move my left hand.
<b>11</b> Il m'a semblé que je ne pouvais pas vraiment dire où se trouvait ma main gauche.	<b>11</b> It seemed like I couldn't really tell where my left hand was.
<b>12</b> Il m'a semblé que ma main gauche avait disparu.	<b>12</b> It seemed like my left hand had disappeared.
<b>13</b> Il m'a semblé que ma main gauche était hors de mon contrôle.	<b>13</b> It seemed like my left hand was out of my control.

In addition to the SRHI-Q, participants answered a series of brief statements to assess their Perceived Similarity to the experimenter. Specifically, we adapted the statements based on the works of Myers & Hodges (2011), Goldstein & Cialdini (2007), and Batson et al. (1997). Participants rated their feelings on a 9-point Likert scale, ranging from “Not at all” to “Extremely,” regarding the extent to which they felt: (1) they shared similar attributes with the experimenter; (2) they had a shared identity with the experimenter; and (3) they were similar to the experimenter. The composite score of the three statements constitute the measure of Perceived Similarity (Goldstein & Cialdini, 2007). Again, it was expected significantly increased scores to each statement following synchronous stroking sessions.

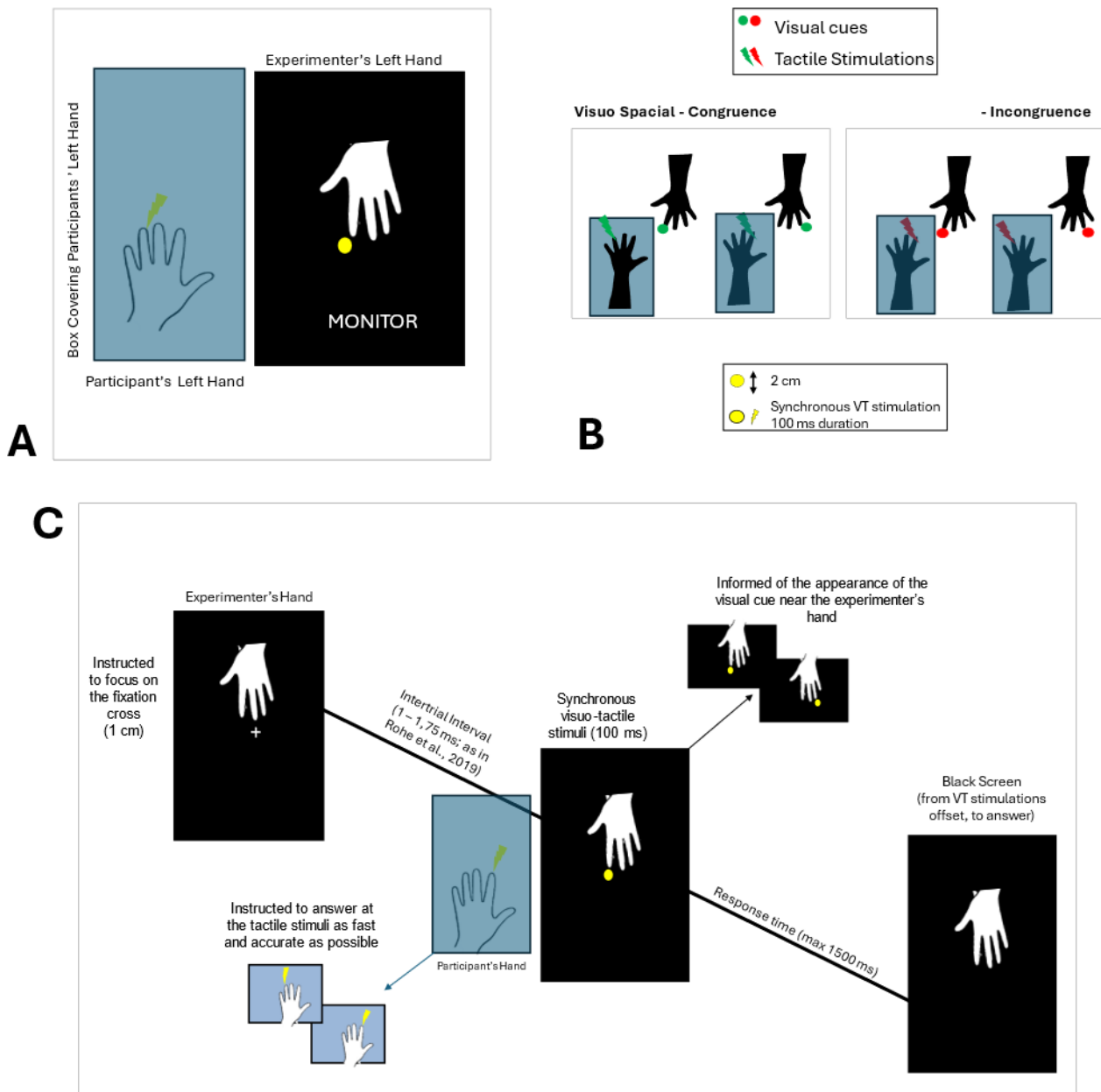
### *Interpersonal Intersensory Task*

Alternatively to the SRHI-Q, participants performed the IIT after one of the two stroking sessions in each block. Again, it is important to note that the SRHI-Q and IIT order of presentation was counterbalanced across participants. Participants were informed that a mechanical stimulator would deliver a tactile stimulus to their hidden left index or ring fingers. Two solenoids were positioned underneath the target fingers, integrated within the platform that balanced the hands' elevation, ensuring no interference with the stroking session and preventing hand movements between stroking sessions (Fig. 4-A). Participants were instructed to accurately discriminate the perceived location of the tactile stimulus (left index or ring finger) as quickly as possible. A keyboard was placed beneath their right hand with a 45-degree angle (Fig 1-B), allowing them to respond by pressing the left arrow key to indicate stimulation on the left ring finger (with their right index finger) or the right arrow key for stimulation on the left index finger (with their right ring finger).

Participants were informed that a task-irrelevant visual cue would appear on the monitor, close to the index and ring fingers of the experimenter's hand. However, they were instructed to respond solely to the tactile stimulation on their own hand. Thus, the tactile stimuli delivered to the participants' fingers, combined with the visual cues near the experimenter's fingers, could result in: (1) spatially congruent visuo-tactile interpersonal stimulations, and (2) spatially incongruent visuo-tactile interpersonal stimulations (Fig. 4-B)(Visuo-Spatial Congruency Factor). Visuospatial congruency refers to the spatial alignment between the stimulated fingers of the participant and the experimenter. From the participant's perspective, when considering the midline of both hands, the congruent condition occurs when the visual cue near the experimenter's hand and the tactile stimulus on the participant's hand are spatially aligned (e.g., both on the left or both on the right). In contrast, in the incongruent condition, the visual and tactile cues are misaligned (e.g., one on the left and the other on the right).

During the task, participants were asked to focus on a 1 cm fixation cross presented 3 cm away from each visual cue, close with the middle finger position of the experimenter. The fixation cross was displayed during an intertrial interval ranging from 1 to 1.75 seconds, as in Rohe et al. (2019). The visual cues presented near the experimenter's fingers measured 2 cm in diameter, consistent with previous studies (in between the dimensions used in Teramoto, 2018; Brozzoli et al., 2013). The experimenter's left index finger was positioned at the same height as the participant's hidden left index finger, with a distance of 25 cm separating the two fingers. The duration of the visuo-

tactile stimulation was 100 ms. Participants were allowed to respond only after the visuo-tactile stimuli offset, to minimize false positive responses. If no discrimination response was recorded within 1,500 ms of the target offset, the trial would terminate; otherwise, the participant's response concluded the trial (similar to Spence et al., 2004). The complete trial timeline is illustrated in Fig. 4-C.



**Figure 4. A) Setup display.** The participant's left hand is shown hidden in the box, while the experimenter's left hand is shown on the monitor. **B) Visuospatial conditions.** Tactile stimulation is applied on participant's left index/ring fingers, while visual cues are showed near the experimenter's left index/ring fingers. Visuo-spatial congruence (left panel) and incongruence (right panel) between the visuo-tactile stimuli is shown. In the congruent condition, the visual and tactile cues match spatially, while in the incongruent condition, they do not. **C) Task timeline.** Participants are instructed first to focus on a fixation cross (1 cm). Synchronous visuo-tactile stimuli (100 ms) are delivered to the participant's and experimenter's fingers, and participants must respond as quickly as possible to the tactile stimuli. Visual cues near

the experimenter's hand are task-irrelevant but can be spatially congruent or incongruent with the tactile stimulation on the participant's hand. Responses must be made within 1500 ms from the stimulus offset.

In each block, participants completed 20 trials per condition (Congruent/Incongruent), resulting in a total of 40 trials per block (Asynchronous/Synchronous), consistent with the prior literature (Maravita et al., 2003; Spence et al., 2004). Additionally, each condition included 6 catch trials without any visual distractor (only tactile stimulation), 12 per block, which were not included in subsequent analyses. The total duration of the task was approximately 2 minutes, with stimuli presentation and randomization managed by E-Prime 3 software (Psychology Software Tools Inc., Pittsburgh, PA).

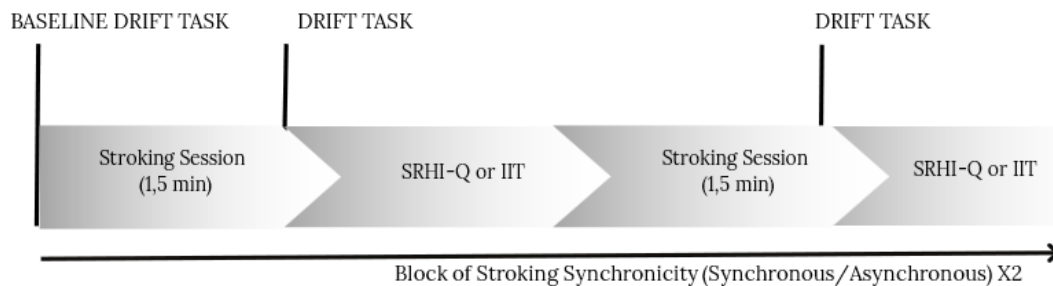
### **Procedure**

Upon signing the informed consent documents and receiving detailed instructions, participants placed their left hand inside the box. They engaged in a brief training session lasting approximately 1 minute, during which they experienced both types of stroking (synchronous and asynchronous). Following this, the experimenter explained the procedure for assessing proprioceptive drift and recorded baseline measurements for the first block. Proprioceptive drift tasks were conducted at baseline before the start of each block to evaluate the impact of stroking on the reported position of the index finger across sessions. Participants also received training for the IIT, completing a trial run consisting of 10 attempts. If participants' accuracy fell below 50% during the trial, the experimenter requested that they repeat the IIT training.

Once instructions and trial runs were completed, the experiment commenced. Two sessions of stroking (lasting 1.5 minutes each) and proprioceptive drift assessments were performed twice per block (synchronous and asynchronous). Following each stroking, participants completed the proprioceptive drift measurements and either the IIT or the SRHI-Q in a counterbalanced order, resulting in one completion of each per block (synchronous/asynchronous). The presentation order of the synchronous and asynchronous blocks was also counterbalanced across participants.

Demographic information and inquiries regarding participants' levels of socio-tactile interaction were collected before the experiment to verify compliance with inclusion criteria. All questionnaires were administered via computer using the Qualtrics platform (Qualtrics International Inc.).

The entire experiment (Fig. 5) lasted, on average, approximately 1.5 hours.



**Figure 5. Experimental Procedure Overview.** Proprioceptive drift was assessed at baseline before each block, and after each subsequent stroking sessions (1.5 minutes each). Participants completed the IIT or SRHI-Q in a counterbalanced order after each drift measurement. The presentation order of synchronous and asynchronous blocks was also counterbalanced across participants. At the end of the experiment, various questionnaires were administered.

### 2.3.3. Data Analyses and Results

#### *Proprioceptive Drift Task*

In our analysis, the origin (zero point) in our drift measurement was established at the corner of the box positioned along the midline between the participant and the experimenter (see Fig. 3), specifically on the side between the experimenter's hand and body. To obtain a unified measure of proprioceptive drift, we calculated an index by computing the Euclidean distance from the origin between reported drift along the x and y axes ( $\sqrt{(x^2+y^2)}$ ), representing the perceived hand position relative to the partner. Data were collected at three-time points: baseline, session 1, and session 2 for each block (synchronous/asynchronous). To measure proprioceptive drift accurately, we subtracted the proprioceptive judgment recorded in each post-stroking manipulation from the baseline measurement obtained before the manipulation. This approach allows us to interpret positive drift values as an increased proprioceptive drift toward the experimenter's hand. No corrections were made for the real hand position when calculating the proprioceptive drift.

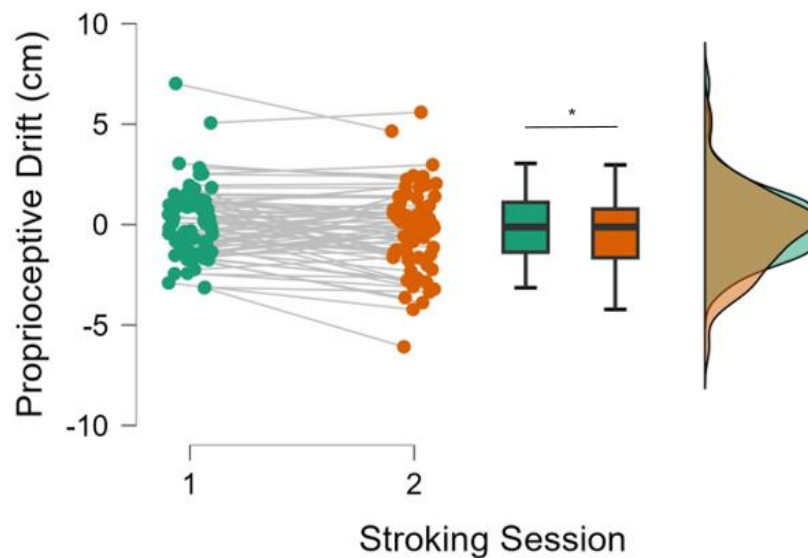
Since the data generally met the assumption of normality (Shapiro-Wilk test  $p > 0.05$ ), a 2 (Stroking Session: 1, 2)  $\times$  2 (Stroking Synchronicity: Synchronous, Asynchronous) within-subject repeated measures Analysis of Variance (ANOVA) was conducted. Holm-corrected pairwise comparisons were performed to explore significant interactions further (Table 2).

**Table 2. Results of repeated Measures Analysis of Variance on Proprioceptive Drift (in cm) with Effect Sizes (partial eta-squared) reported.** Asterisks flags significant p-values.

Cases	Sum of Squares	df	Mean Square	F	p	$\eta^2_p$
Stroking Synchronicity	0.745	1	0.745	0.120	0.731	0.003
Residuals	217.427	35	6.212			
Stroking Session	5.298	1	5.298	4.719	0.037*	0.119
Residuals	39.295	35	1.123			
Stroking Synchronicity * Stroking Session	0.013	1	0.013	0.015	0.902	4.349×10 <sup>-4</sup>
Residuals	30.398	35	0.869			

Note. Type III Sum of Squares

The analyses revealed a main effect of Stroking Session ( $F(1,35) = 4.72$ ;  $p = 0.04$ ;  $\eta^2_p = 0.12$ ), but no main effect of Stroking Synchronicity ( $p = 0.73$ ) or a significant interaction ( $p = 0.90$ ). Specifically, Session 2 shows a significant drift toward the experimenter's ( $M = 0.05$ ;  $SD = 0.48$ ) than Session 1 ( $M = 0.03$ ;  $SD = 0.50$ ) (Fig. 7).



**Figure 7. Main effect of Stroking Session on Proprioceptive Drift (in cm).** A significant main effect of Stroking Session was observed ( $p = 0.04$ ), indicating more significant drift in Session 2 compared to Session 1. Error bars represent the standard deviation.

#### *'Social' Rubber Hand Illusion Questionnaire*

According to Shapiro-Wilk test, not all conditions met the normality criteria ( $p < 0.05$ ), therefore a 3 Component (Ownership, Location, Hand Loss)  $\times$  2 Stroking Synchronicity (Synchronous, Asynchronous) Friedman ANOVA was conducted, that resulted significant ( $\chi^2 = 96.04$ ;  $p = 0.00$ ).

Wilcoxon Matched Pairs Test were performed to follow up on comparisons in each Component between the Synchronicity levels, corrected for multiple comparisons (e.g., 2 in our case) to mitigate the risk of false positives (corrected significance level of  $p < 0.025$ )(Table 3).

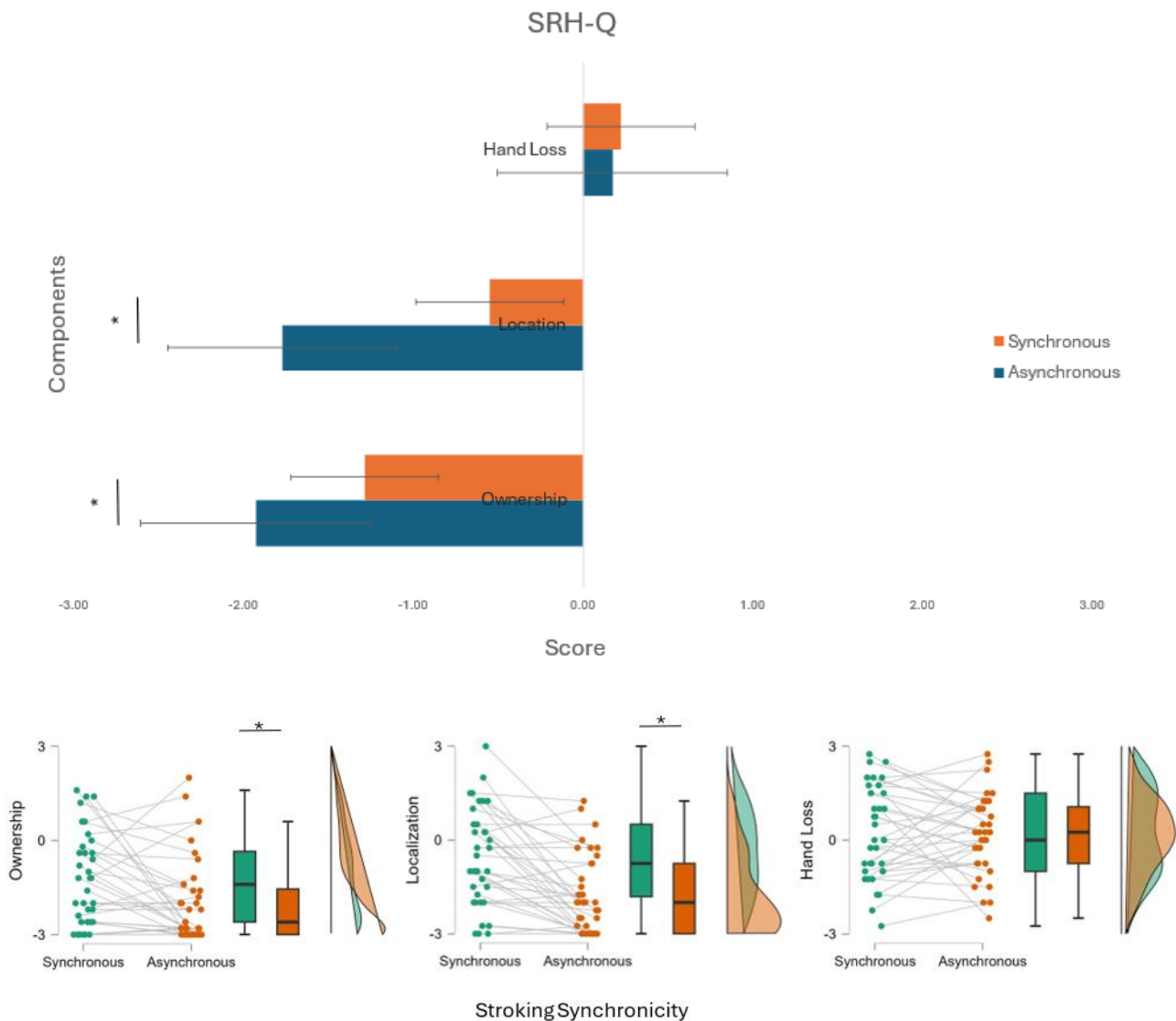
**Table 3. Wilcoxon Matched Pairs Test.** Asterisks flag significant p-values ( $< 0.025$ ).

			W	z	p	Rank-Biserial Correlation	SE Rank-Biserial Correlation
Ownership, Asynchronous	-	Ownership, Synchronous	112.50 0	- 2.270	0.024*	-0.483	0.209
Location, Asynchronous	-	Location, Synchronous	38.000	- 4.115	$< .001^*$	-0.847	0.203
Hand Loss, Asynchronous	-	Hand Loss, Synchronous	299.50 0	- 0.254	0.805	-0.049	0.191

*Note.* Wilcoxon signed-rank test.

The test highlight that Ownership was significantly experienced after the Synchronous sessions ( $M = -1.28$ ;  $SD = 1.48$ ) compared to the Asynchronous ones ( $M = -1.93$ ;  $SD = 1.38$ ) ( $p = 0.02$ ). Furthermore, for the Location component, in the Asynchronous strokings ( $M = -1.77$ ,  $SD = 1.26$ ), participants reported less confusion between the location of their hand and the observed one, but also regarding the performed and received touch, compared to the Synchronous one ( $M = -0.55$ ,  $SD = 1.57$ ;  $p < 0.00$ ). These differences between Stoking Synchronicities (Synchronous, Asynchronous) were absent in the Hand Loss component ( $p = 0.81$ ) (Table 3).

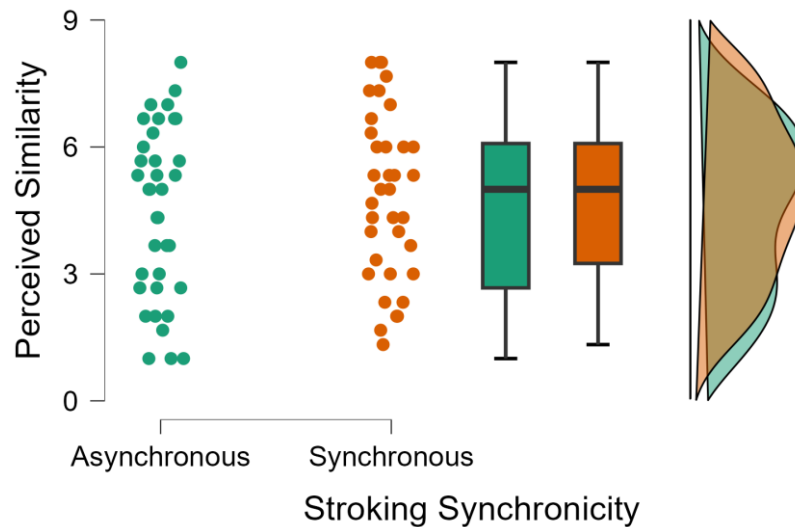
In summary, the SRHI-Q Components of Ownership and Location were affected by the Stroking Synchronicity, and the result indicates that the illusion was less experienced after the Asynchronous stroking sessions, in line with previous literature (Longo et al., 2008; Tsakiris et al., 2010). In contrast, the Hand Loss component was overall reported after both the Synchronous ( $M = 0.22$ ,  $SD = 1.48$ ) and Asynchronous stroking sessions ( $M = 0.17$ ,  $SD = 1.27$ )(Fig. 8).



**Figure 8. Results of the interaction between Stroking Synchronicity (Synchronous, Asynchronous) \* SRHI-Q Components (Ownership, Location, Hand Loss).** Asterisks flag significant interactions. Error bars represent the standard error.

### *Perceived Similarity*

According to the Shapiro-Wilk test data met normality criteria ( $p > 0.05$ ), than an ANOVA was conducted on the effect of the Stroking Synchronicity (Synchronous, Asynchronous) on the composite scores to the three statements (general perceived similarity, shared dentity and attributes) of the Perceived Similarity measure (see *SRHI-Questionnaire* section). The test resulted in no significant difference in the perceived similarity with the experimenter across conditions and overall moderately positive scores ( $M = 4.66$ ;  $SD = 1.99$ ) (Figure 9).



**Figure 9. Results of the Anova on Perceived Similarity of the Stroking Synchronicity.** Error bars represent the standard error.

### Interpersonal Intersensory Task

The dependent variables employed for the statistical analysis comprises the Reaction Times (RTs) for the Tactile Response (in ms), corresponding to the offset timing between the appearance of the visuo-tactile stimulation and the discrimination response of the participant to the tactile stimulus, measuring VTI effects in the paradigm. Accuracy (ACC) was also considered, corresponding to the correct discrimination of the tactile stimulation location (index/ring).

We refined the dataset by excluding catch-trials. From the RTs, also inaccurate trials were excluded from the analyses and trials where participants missed the tactile stimulation (i.e., did not initiate keyboard response). Furthermore, to ensure data quality and avoid extreme values that could have been linked to attentional biases or fatigue errors, at the individual trial level, any value exceeding or falling below  $\pm 2.5$  standard deviations of the individual mean for each experimental condition was treated as outlier and excluded from the analyses (Ratcliff, 1993). Additionally, two participants could not perform the task, while one participant exhibiting RTs that diverged by  $\pm 2.5$  standard deviations from the group mean was identified as an outlier (Ratcliff, 1993) and the three were excluded from all the subsequent analyses.

#### *Reaction Times*

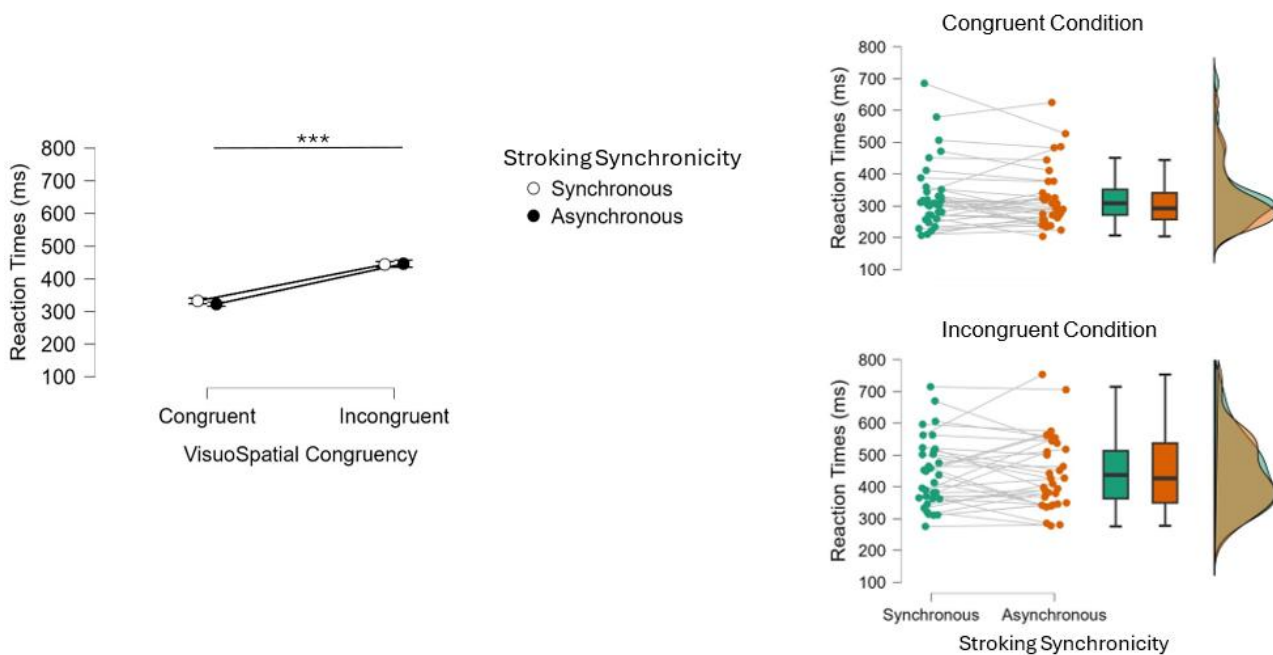
Since the data significantly deviated from normality in the Congruent conditions according to the Shapiro-Wilk test ( $p < 0.00$ ), a 2 Stroking Synchronicity (Synchronous, Asynchronous) x 2

VisuoSpatial Congruency (Congruent, Incongruent) Friedman ANOVA was conducted on RTs, that resulted significant ( $x^2 = 68.71$ ;  $p = 0.00$ ). Thus, Wilcoxon Matched Pairs Test were performed (Table 4), corrected for multiple comparisons to mitigate the risk of false positives (e.g., 4 in our case), resulting in a corrected significance level of  $p < 0.01$ .

**Table 4. Wilcoxon Matched Pairs Test on RTs.** Asterisks indicates tests are significant at  $p < .01$

Pair of Variables	Valid	T	Z	p-value
Synchronous_Congruent * Synchronous_Incongruent	33	1.000000	4.994058	0.00*
Synchronous_Congruent * Asynchronous_Congruent	33	218.0000	1.116739	0.26
Synchronous_Incongruent * Asynchronous_Incongruent	33	272.0000	0.151877	0.88
Asynchronous_Congruent * Asynchronous_Incongruent	33	0.00	5.011926	0.00*

These results suggest a main effect of VisuoSpatial Congruency (all  $p < 0.00$ ), reflecting the classical VTI effects with slower RTs for Incongruent trials ( $M = 445.02$ ,  $SD = 112.11$ ) compared to Congruent ones ( $M = 327.41$ ,  $SD = 100.72$ ) (Maravita et al., 2003; 2002; Spence et al., 2000; 2004). While, the interactions between the factor levels did not achieve statistical significance (all  $p > 0.05$ ), showing that synchronous and asynchronous stroking did not differ in the impact on the VTI effect in the IIT task (Fig. 10).



**Figure 10. IIT - Reaction Time (in ms) Results.** The graphs show the data distribution in the Interpersonal Intersensory Task of the RTs for Tactile Responses. The classical VTI effect was found for Congruent\*Incongruent trials in both Synchronous and Asynchronous sessions.

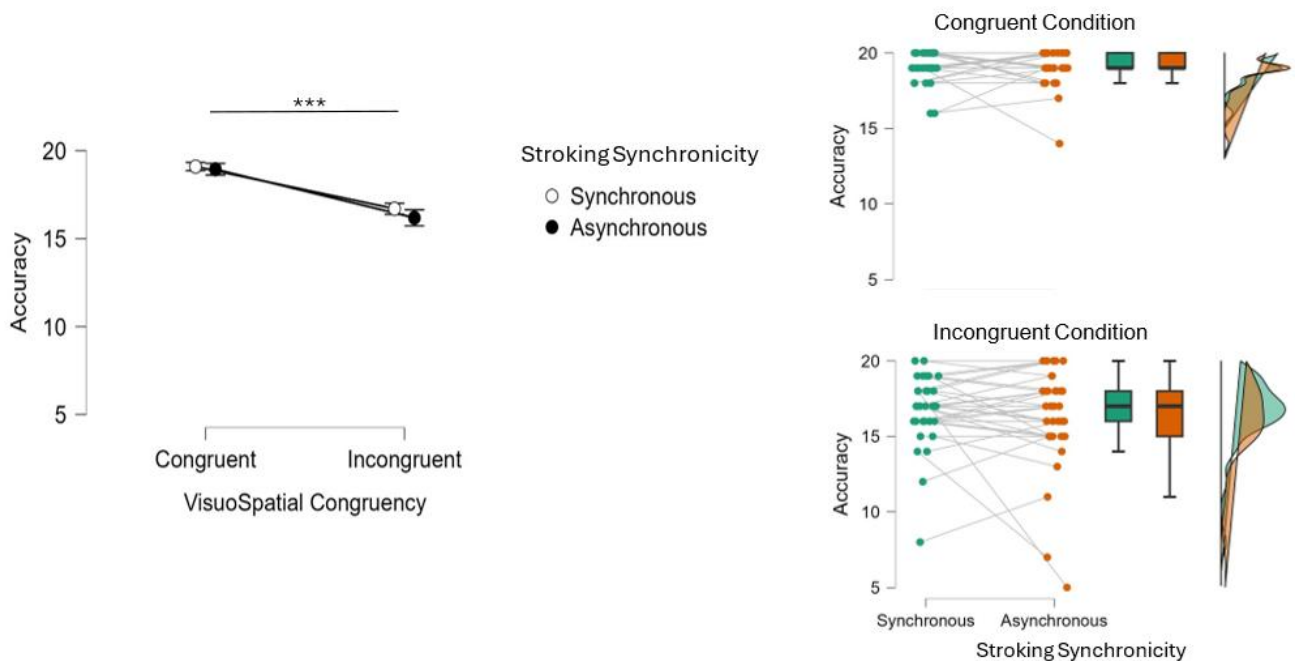
*Accuracy*

A 2 Stroking Synchronicity (Synchronous, Asynchronous) x 2 VisuoSpatial Congruency (Congruent, Incongruent) Friedman ANOVA was conducted on ACC (20 trial per condition) and resulted significant ( $\chi^2 = 43.11$ ;  $p = 0.00$ ). Wilcoxon Matched Pairs Test were performed, corrected for multiple comparisons to mitigate the risk of false positives (e.g., 4 in our case), resulting in a corrected significance level of  $p < 0.01$ . The results complemented the one found for RTs (Table 5).

**Table 5. Wilcoxon Matched Pairs Test on ACC.** Asterisks indicates tests are significant at  $p < .01$

Pair of Variables	Valid	T	Z	p-value
Synchronous_Congruent * Synchronous_Incongruent	27	3.500000	4.456637	0.00*
Synchronous_Congruent * Asynchronous_Congruent	22	113.0000	0.438286	0.66
Synchronous_Incongruent * Asynchronous_Incongruent	29	209.0000	0.183797	0.85
Asynchronous_Congruent * Asynchronous_Incongruent	28	19.50000	4.178556	0.00*

The results suggested a main effect of VisuoSpatial Congruency (all  $p < 0.00$ ) reflecting the classical interferent effects in reduced ACC for Incongruent trials ( $M = 16.44$   $SD = 2.92$ ) compared with Congruent ones ( $M = 19.02$ ,  $SD = 1.10$ ). The interactions between factor levels did not achieve statistical significance (all  $p > 0.05$ ), showing that synchronous and asynchronous stroking sessions did not impact the accuracy of the VTI effect in the IIT task (Fig. 11).



**Figure 10. IIT - Accuracy (across 20 trials) Results.** The graphs show the data distribution in the Interpersonal Intersensory Task of the ACC for Tactile Responses. The classical VTI effect was found for Congruent\*Incongruent trials in both Synchronous and Asynchronous stroking sessions.

### 2.3.4. Discussion

The present findings highlight the intricate interplay between multisensory visuo-tactile integration and body ownership in social contexts. Here, we presented a novel approach that goes toward a ‘social’ rubber hand illusion task, revealing the impact that interpersonal exchanges can have on body and space perception.

The proprioceptive drift task revealed a significant effect of session repetition, through stroking synchronicity had no impact on participants’ reports of their hand position relative to the experimenter’s hand. These results indicate that repeated exposure to the task increased proprioceptive drift, irrespective of whether the stroking was synchronous or asynchronous. This finding contrasts with traditional RHI studies, which emphasize the crucial role of synchronicity in inducing proprioceptive changes (Botvinick & Cohen, 1998; Tsakiris et al., 2005, 2006). A potential explanation for the absence of a synchronicity effect could be the social nature of our paradigm, where the proximity and the mutual interaction with another individual might already be

sufficient to induce proprioceptive adjustments toward the partner, reducing the influence of stroking synchronicity.

These results align with previous research stating that embodiment sensations and shifts in proprioception can be dissociated (Rohde et al., 2011; Gallagher et al., 2021). Notably, the SRHI-Q revealed a significant effect of stroking synchronicity on the Ownership and Location components, reinforcing the importance of synchronous multisensory signals in evoking a sense of embodiment, consistently with previous literature (Botvinick & Cohen, 1998; Longo et al., 2008; Tsakiris et al., 2010). In the synchronous condition, participants reported a stronger sense of ownership and confusion about the location of their hand, as well as perceived congruence between received and executed touch. Interestingly, this effect showed even for a hand in an implausible position, as expected. This finding stands in contrast to earlier studies that reported the RHI failing to manifest for unrealistic hands orientations (Farnè et al., 2000; Tsakiris & Haggard, 2005). A potential explanation lies in the social mutual nature of our task, which may have created a more compelling multisensory integration in synchronous conditions so that the sense of embodiment extended to the partner as a result of integrated afferent and efferent inputs. The absence of a significant synchronicity effect on the Hand Loss component further supports this interpretation, suggesting that this novel paradigm might induce a robust sensation of touching one's own hand while stroking the partner.

The lack of effect on Perceived Similarity indicates that the illusion might not extend to altering higher-order interpersonal perceptions, such as the sense of shared identity or personal attributes. Alternatively, the interaction dynamics in our task may have fostered a general perceived similarity with the partner as a result of the mutual engagement in both synchronicity conditions. This interpretation complements previous findings in interpersonal motor task, which are proven to increase other social processes such as affiliation (Hove & Risen, 2009), trust (Wiltermuth & Heath, 2009) and conformity (Wiltermuth, 2012).

Lastly, the IIT results observed are consistent with the classical visuo-tactile interference effects (Spence et al., 2000; Maravita et al., 2003). However, these effects did not vary with the prior synchronous or asynchronous manipulations, suggesting that multisensory integration mechanisms underlying VTI may be distinct from those contributing to the subjective experience of the RHI. These findings, however, align with Fossataro et al. (2023), who proposed that the representation of self-space may shrink relative to another person during tactile tasks, which could explain why tactile discrimination performance remained stable across conditions despite subjective changes in body ownership.

## **Limitations & Future Directions**

While this study provides novel insights into the social dimensions of body ownership and perception, it is important to acknowledge several limitations.

Firstly, the sample consisted exclusively of same-gender female dyads. Although this choice helped control potential confounding factors, such as differences in hand appearance or gender dynamics, it limits the generalizability of the findings. Future research should investigate mixed-gender and male dyads to assess whether the effects observed in this study hold across different gender and populations. This could be addressed in future studies by using video-recordings or kinematic tracking improving the control over this factor.

Secondly, while the mutual movement set-up introduced here guarantee an ecologically valid interaction between participants and the experimenter, differences in the precision, intensity, and timing of the stroking should have been controlled better, to ensure the absence of variability across participants.

Thirdly, the design of the asynchronous condition, meant to disrupt predictability and visuo-tactile integration, was not aligned with the traditional RHI methods. The disruption was not limited to timing and synchrony but extended to include differences in movement type and location, potentially introducing additional dynamics beyond the intended disruption of multisensory integration. Similarly, the measurement of the proprioceptive drift differed from previous studies, potentially affecting the results. While these modifications were implemented to accommodate the novel aspects of the paradigm introduced in this study, future research should aim to more closely align the asynchronous condition and the proprioceptive drift task with the established classical RHI protocols and our group is already structuring a follow-up study considering all these aspects.

## **Conclusions**

In conclusion, while stroking synchronicity enhances feelings of body ownership and alters subjective body perceptions, it appears to have minimal effects on higher-order social cognition or automatic VTI processes. These findings suggest a dissociation between the multisensory mechanisms that govern body ownership and those that govern sensory processing in interpersonal contexts. Future research should deepen the understanding on how these processes and social factors interact.

This study adds to the literature on body and space representation, showing that synchronicity in interactive contexts fosters embodiment, even with an implausible hand position, though its impact on social cognition and visuo-tactile integration is limited. While these findings are primarily based on subjective measures, this study advances our understanding multisensory integration by extending it to social contexts and calls for future exploration of the underlying neural mechanisms.

## **2.4. The Hidden Cost of Outgroup Coordination: Action Delays in Joint Actions.**

### **2.4.1. Introduction**

Racial differences are a straightforward cue to rapidly categorize others within a stereotype (Cosmides et al., 2003). This immediate categorization can shape intergroup interactions, leading to implicit biases and prejudicial attitudes, resulting in discrimination and social inequality (Devine, 1989; Amodio, 2008; Maister et al., 2013). Examples of biases in favor of one's racial group, compared to racial out-groups, have been documented in a series of domains spanning from judicial decision-making to sport, also including medical practice (Do Bù et al., 2023; Shayo & Zussman 2011, Zhang, 2017). Although the study of racial biases has a long tradition in social psychology (Dovidio et al., 2005), neuroscientists have started to systematically investigate the neural bases of racialized cognition only in the last decades (Ito & Bartholow, 2009; Kubota, Banaji & Phelps, 2012). Some authors have focused on the neural processing of faces and bodies belonging to racial in-groups and out-groups (e.g., Ito & Urland, 2005; Watson & De Gelder, 2017; Vizioli et al., 2010; Zhou et al, 2020). Differently, other researchers have shown more interest in studying mirroring and empathy responses in the brain when observing individuals sharing the same or a different ethnicity compared to the participant (e.g., Avenanti, Sirigu & Agliot, 2010; Azevedo et al., 2013; Contreras-Huerta et al., 2013; Gutsell & Inzlicht ,2010; Sheng & Han, 2012). The latter showed how racial biases affect neural responses to observed actions and pain. For instance, Gutsell and Inzlicht (2010) found motor cortex activation when participants watched actions performed by their in-group members, but reduced activation for racial out-groups. Another relevant work was conducted by Avenanti et al. (2010), who found that observing pain in same-race individuals triggered typical neural empathy responses in the motor and sensorimotor cortex, while this response was absent for other races. This lack of neural resonance correlated with implicit racial biases measured by the Implicit Association Task (IAT; Greenwald et al., 1998). Han (2018) provided an extensive summary of research evidence concerning differential brain activations when witnessing same-race or other-race individuals experience pain. His review highlighted how the neural empathetic response is stronger when elicited by a person from the same ethnic group than a member from an ethnic outgroup, which seems to be shaped by sociocultural and environmental factors.

Nevertheless, the author also underlined that this biased neural response can be diminished by interventions that reduce the distance between the in-group and the out-group.

Hence, racial group membership and stigma can impact mirroring and empathic processes (Decety et al., 2009; Stürmer et al., 2006; Yabar et al., 2006; Link & Phelan, 2001), but those effects may be mitigated. To better understand how to disrupt the bias, researchers started investigating the neural underpinnings of racial prejudice and discrimination in the first place. It was suggested that reduced medio-prefrontal cortex (mPFC) activity may reflect a dehumanization process characterized by decreased empathy toward others (Amodio, 2014). Indeed, previous literature indicates that the mPFC is implicated in racial stereotyping (Interactiveman et al., 2010; Gilbert et al., 2012). The mPFC appears to be significantly engaged when individuals regulate their intergroup responses in accordance with non-prejudiced norms, suggesting a crucial self-regulatory role for this area in mitigating prejudice (Amodio et al., 2006). This involvement of the mPFC is attributed to its position at the intersection of neural regions associated with the representation of one's actions (dACC) and the representation of another person's actual and anticipated actions (anterior mPFC and orbital frontal cortex) (Amodio & Frith, 2006).

Taken together, these findings indicate how racial group membership could modulate the activity of brain areas involved in the simulation and anticipation of others' actions, potentially affecting cooperation and coordination in social interactions. It remains an open question whether implicit racial attitudes and, subsequently, the outcomes of social interactions, may be altered by experimentally inhibiting the neural areas involved in racial stereotyping.

Understanding the neural basis of social interaction with the out-group could provide valuable insights into social cognition processes and into developing interventions aimed at reducing prejudice and racial biases. To do so, the present study combines a joint action task and the use of repetitive transcranial magnetic stimulation (rTMS) to explore the complex role of the mPFC in regulating interaction with an out-group member to achieve a common goal. In three different experiments, participants were presented with two virtual partners during the joint action task: an out-group (Black) or an in-group (White) avatar (the control condition). The instruction was to synchronize with the partner to gasp together a bottle-shaped object in front of them, performing the same or opposite movement relative to the avatar. A recent study has already proven how a partner's racial membership can modulate interactive behaviors (Sacheli et al., 2015a). Sacheli et al. (2015a), showed how belonging to different ethnicities reduces predictive embodied simulations of an out-group partner's movements. Therefore, before the joint action task, a continuous theta burst

stimulation (cTBS; Huang et al., 2005) was administered to investigate the role of mPFC in coordinating with an out-group partner after a virtual inhibition of the area.

We specifically aimed to investigate whether mPFC inhibition reduces predictive simulation and, consequently, the ability to synchronize during joint action with an out-group member while exploring the previously found modulation of partner racial membership on interaction for shared goals. Furthermore, we aimed to evaluate whether the stimulation could increase explicit and implicit racial biases as measured by the Affect Misattribution Procedure (AMP).

## **2.4.2. Materials and Methods**

### **Participants**

The sample comprised sixty caucasian participants, but 5 participants were excluded from all the analyses since they did not meet our inclusion criteria. Hence, the final sample comprised fifty-six participants (Experiment 1, N = 16; Experiment 2, N = 15; Experiment 3, N = 24; 28 males; mean age =  $24.69 \pm 3.97$ ). All participants were right-handed, as evaluated by the Standard Handedness Inventory (Briggs & Nebes, 1975), and presented normal or corrected-to-normal vision. To be included in the sample, participants had to read and fill out a screening questionnaire that evaluated their suitability to undergo TMS (Rossini et al., 2015). Participants who reported any of the contraindications indicated by the questionnaire and those who reported using drugs with effects on the nervous system or being affected by psychiatric or neurological syndromes were excluded from all three experiments.

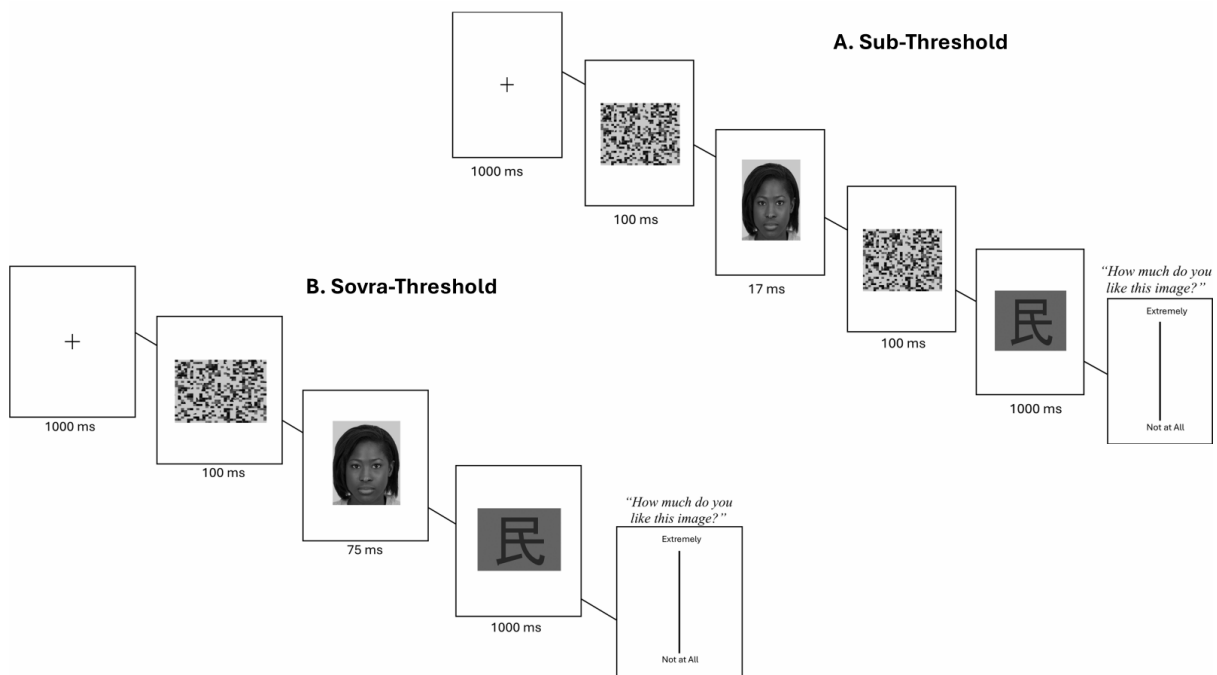
The experimental protocol was approved by the ethics committee located in Fondazione Santa Lucia and was conducted in agreement with the 1964 Declaration of Helsinki. Participants were recruited by email and word of mouth, and written informed consent was signed. No monetary compensation was provided for the participation in the experiment.

### **Affect Misattribution Procedure**

Two versions of the affect misattribution procedure (AMP; Payne et al., 2005), adapted from those used by Boukarras and colleagues (2020), were implemented to indirectly measure the prejudice toward caucasian (in-group) and African (out-group) ethnicities. The AMP is a task widely used to evaluate implicit biases toward specific categories, which requires individuals to evaluate an ambiguous target stimulus (e.g., Chinese characters for people who cannot read Chinese) preceded

by a prime. Given the ambiguity of the target, the evaluation is influenced by the affect automatically elicited by the priming stimulus, even when the participant is overtly asked to ignore it (Payne & Lundberg, 2014). Thus, a prime that evokes a negative affective state in the participant induces a worse judgment of the target.

In the three experiments, participants were asked to evaluate several Chinese ideograms aesthetically, ignoring the stimuli that preceded their appearance on the screen. Those priming stimuli consisted of 10 black and 10 white faces of real individuals selected from the Chicago Face Database (Ma, Correll & Wittenbrink, 2015). A validation of the stimuli was conducted in a sample of caucasian Italian participants ( $N = 30$ ), in which no significant differences emerged in participants' evaluation of white and black faces along a series of dimensions (i.e., age, fearfulness, anger, attractiveness, being baby-faced, disgust, dominance, happiness, femininity, sadness, being threatening, trustworthiness; all  $ps > .05$ ). A first version of the AMP was intended to measure the bias induced by a prime presented below the perceptual awareness threshold (Sub-threshold; Boukarras et al., 2020; Murphy & Zajonc, 1993). In this version, the trial started with a fixation cross at the center of the screen, which was replaced after 1000 ms by a forward mask lasting 100 ms. Then, the image of a white or black person's face (e.g., the priming stimulus) appeared. The face remained for 17 ms and was followed by a scrambled version of the same image (i.e., backward mask). After another 100 ms, a Chinese ideogram (i.e., target stimulus) was shown for 1000 ms. Finally, a vertical Visual Analog Scale (VAS) appeared on the screen, with a label over its top end stating "Extremely" and one below its bottom end stating "Not at all". Participants were asked "How much do you like this image?" and were instructed to click on the segment of the VAS that reflected their judgment (Fig. 1-A). Their answer to the VAS would start the subsequent trial. In a second version of the task, created to investigate the bias produced by a prime shown over the perceptual threshold (Boukarras et al., 2020; Payne et al., 2005), everything was the same except for the absence of a forward mask and the presentation time of the priming faces, which was 75 ms (Sovra-threshold) (Fig. 1-B).



**Fig 1. A) Sub-Threshold Priming:** Participants viewed a fixation cross (1000 ms), followed by a forward mask (100 ms). A face prime (Black or White) was presented briefly (17 ms), immediately followed by a backward mask (100 ms). Afterward, a Chinese ideogram appeared for 1000 ms. Participants then rated how much they liked the ideogram using a Visual Analog Scale (VAS) from "Not at all" to "Extremely". **B) Sovra-Threshold Priming:** This version followed the same procedure as panel A, but without a forward mask, and the face prime was presented for a longer duration (75 ms), making it consciously visible. Participants again rated the ideogram using the VAS.

Notably, the pairing and the order of presentation of priming faces ( $N = 20$ ) and ideograms ( $N = 36$ ) were randomized through E-Prime2 software (Psychology Software Tools Inc., Candidi et al., 2017; Era et al., 2023; Era, Aglioti & Candidi, 2020; Moreau et al., 2020; Pittsburgh, PA). Each ideogram was presented twice, once after a caucasian face and once after an African face. Moreover, the size of the prime and the scrambled images was the same (362 px in length, 450 px in height), while the Chinese ideograms were bigger (512 px large, 344 px in height). The total number of trials in a block was 36 (equal in both versions), in both tasks, there were two blocks, for a total of 72 trials per task. Participants could also take a break when needed after the first block of trials in each task.

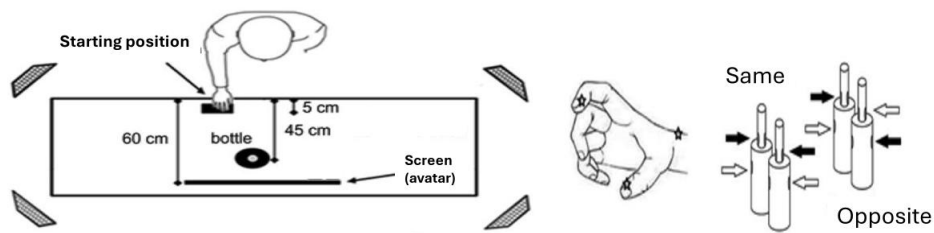
### Experimental set-up and joint grasping task

To study participants' motor coordination with virtual partners representing different ethnic groups, we adopted a joint grasping task, which has been repeatedly and successfully applied to the study of the neural underpinnings of dyadic motor performance (Boukarras et al., 2021; Sacheli et al., 2015a, Sacheli et al., 2018). The task required participants to coordinate with two virtual partners (one at a

time) to synchronously grasp a bottle-shaped object. To reach their goal, participants had to continuously track and adapt to their partner's movements, an avatar. This mechanism underlines how the task could reproduce human-human joint actions in a more controlled set-up (Sacheli et al., 2012; 2013; Candidi et al., 2015; Era et al., 2018; Boukarras et al., 2021).

Participants were seated at a rectangular table (120 x 75 cm), a LCD monitor with a 1024 x 768 resolution was placed about 60 cm from them, while a bottle-shaped object composed of two superimposed cylinders (37 cm height) was placed between the monitor and participants, at a distance of 45 cm from the eyes of the latter. To grasp the object, participants were instructed to touch the front side with their right thumb and the backside with their right index. The bottle consisted of two stacked rectangles with different widths. The grasp could be performed on the upper or lower half of the object. Participants were instructed to use a precision grip on the upper section (2.7 cm diameter) or a power grip on the lower section (6.5 cm diameter). Four plates sensitive to touch were located on the target object, two on the upper half and two on the lower half, at distances of 15cm and 22cm along the vertical height of the object so that the exact time of participants' grasp was recorded.

Following, we report the timeline of a single trial (Fig. 3-A). Before the trial started, participants placed their right thumb and index fingers, joined, on a button located on their right, at a 34 cm distance from the bottle-shaped object. Subsequently, a fixation cross appeared at the bottom left of the screen for 300 ms, corresponding to the location where the partner's hand would be displayed and the partner's movement started. After the cross disappeared, a clip showing the partner moving his right arm to grasp a virtual bottle-shaped object identical to the one on the table was displayed on the monitor. Before the clip was reproduced, an audio instruction lasting around 193ms was transmitted via headphones, which consisted of one of two words: "same" or "opposite". The "same" instruction informed participants to imitate the movement performed by the partner, while the "opposite" instruction indicated they had to perform a complementary movement. At this point, participants lifted their fingers from the starting button and tried to grasp the object as synchronously as possible with their partner. The virtual partner could grasp the target object on its lower or upper half. Therefore, an imitative movement (Same) implied that the participant and the partner grasped the object on the same half (i.e., both on the lower or both on the upper half). Conversely, a complementary movement (Opposite) required participants to grasp the object on the opposite half compared to the partner (i.e., participant on the lower half and virtual partner on the upper half or vice versa).



**Fig 2. Set-Up Illustration.** The left panel displays the table setup, which includes the participant's starting position, as well as the distances between the participant, the avatar, and the bottle-shaped object to be grasped. The right panel highlights the fingers involved in the grasping action (the index finger and thumb) and shows the possible movement types in relation to the partner (either opposite or same).

Participants were unaware of the ending location of the partner's movement before a trial started. Thus, they had to monitor their movement online to grasp the target object on the correct half. A photodiode was positioned on the lower right of the monitor to measure the exact time of the virtual partner touch, avoiding potential errors due to the refresh rate. When the partner grasped its object, a white rectangle appeared on the portion of the screen where the photodiode was located. The change in luminosity caused by the rectangle was recorded by the photodiode, which communicated it to a trigger station (BrainTrends Ltd., Italy), which in turn sent a signal to the E-Prime2 software (Psychology Software Tools Inc., Pittsburgh, PA), used for the timings acquisition and trial administration.

### Virtual interaction partners

The two male virtual partners used in the joint grasping task were created following the same procedure used by previous studies (Candidi et al., 2017; Moreau et al., 2020). The kinematics of a real individual performing a series of movements to grasp the bottle-shaped object with their right hand were recorded by means of a SMART-D motion capture system [Bioengineering Technology & Systems (B|T|S)]. Following, those kinematics were implemented on two avatars through MotionBuilder 2017 (Autodesk, Inc.), while the 3DS Max 2017 software (Autodesk, Inc.) was used to implement the rest of the virtual scenario in which the avatars were shown. The skin color of the avatars was modulated to represent two partners from different ethnic groups (one caucasian and one African). Moreover, only part of the avatars' body was displayed, from their shoulders to the end of their torso. Their face was not shown to clear out any influence related to trustworthiness, dominance, and other dimensions that people automatically infer just by looking at faces with no prior information (Oosterhof & Todorov, 2008; Todorov et al., 2008, 2013). A series of clips showing the virtual partners grasping the bottle-shaped object were implemented using E-Prime2.

Five clips were created for both avatars. In half of the trials, the avatars grasped the object on the lower part (power grip), and in the other half, they grasped it on the upper part (precision grasp). To reduce the predictability of the virtual partners, the duration of their grasping movements was varied between trials. It must be noted that 12.5% of trials contained a movement correction, that is, the partner started to move toward one half of the object and then suddenly changed direction, grasping the other half. This type of correction was introduced to reduce the predictability of the partners' movements and increase participants' need to monitor and adapt online to the partner. Those trials were considered "catch" trials and were not included in the statistical analyses. All video stimuli had a duration of approximately 2000ms.

### **Transcranial magnetic stimulation**

TMS is a pain-Interactive and safe technique that allows the direct stimulation of a small neuronal population (Pascual-Leone et al., 2000). Repetitive TMS (rTMS) is a magnetic stimulation capable of inducing physiological and behavioral changes over a longer time than the stimulation itself (Siebner & Rothwell 2003). rTMS is a technique that only presents contraindications in particular clinical conditions. Therefore, participants' suitability in the study was assessed by employing an ad-hoc questionnaire (Rossini et al., 2015). A 20-second continuous theta burst stimulation (cTBS) paradigm, that has been shown to have an inhibitory effect on the stimulated site starting 5 minutes after stimulation and lasting up to 20 minutes after stimulation (Huang et al., 2005). The procedure was the same as Huang et al. (2005): trains of three 50 Hz pulses were delivered every 200 ms (i.e., at 5 Hz) for 20s (300 pulses in total).

The TMS protocol applied differed between the three studies. As in prior studies targeting the parietal cortex (Rosenthal et al., 2009; Yazar et al., 2017), cTBS was delivered at 80% of the resting motor threshold in experiments 1 and 2 (average intensity:  $45.45\% \pm 6.79\%$  of stimulator output). In experiment 3, the stimulation was applied at 40% of the maximum stimulator output (as in Amoruso et al., 2018; Era et al., 2021), a fixed intensity chosen for the coil properties and output capabilities, but also its safety, since it corresponds to the lowest motor threshold (Wassermann, 1998). Following the stimulation, participants rested for 5 minutes with their right arm in a relaxed position before starting the interactive task.

The cTBS procedure was performed on each site in separate 60-minute sessions (Huang et al. 2005). The session order was counterbalanced across subjects. The procedure in each session involved stimulation of one of the sites (mPFC/Vertex in Experiment 1, Vertex/vPMC in Experiment 2) and the execution of two experimental blocks of joint action, counterbalanced within

and across subjects. In each joint action block, the subjects had to synchronize their movements with the white or black avatar. The performance of these two blocks took 15 minutes, within the time window in which cTBS is effective in modulating cortical excitability (Huang et al., 2005) as in studies similar to those presented here (Sacheli et al., 2015a; Era et al., 2018).

TMS was performed using a 70 mm figure-of-eight coil connected to a transcranial magnetic stimulator (The Magstim Company). Stimulation sites were identified on each participant's scalp using the SofTaxic Navigator system (EMS). The skull reference points (nasion, inion, and two preauricular points) and 61 additional points provide a uniform representation of the scalp digitized with a Polaris Vicra optical tracking system (NDI). The SofTaxic Navigator automatically estimated the coordinates in Talairach space (Talairach & Tournoux, 1988) from a stereotactic template constructed with magnetic resonance imaging (MRI) using an individualized probabilistic head model. Across the three experiments, the system identified and stored the sites that optimally targeted the Vertex for each participant according to the coordinates reported by Okamoto et al. (2004;  $x = 0$ ,  $y = -17$ ,  $z = 63$ ). In experiments 1 and 3, the same procedure was adopted for the mPFC (Mackey & Petrides, 2014;  $x = 5$ ,  $y = 43$ ,  $z = 30$ ), and in experiment 2 for the vPM ( $x = -52$ ,  $y = 10$ ,  $z = 24$ ; Avenanti et al., 2013) as an active control (as in Era et al., 2021). This individualized template preserves the anatomical correlates of the scalp and brain of an average MRI template specific to the participant under investigation. The resulting mean stimulation coordinates were cTBS was applied in experiments 1 and 3 were  $x = 6.15 \pm 1.98$ ,  $y = 48.93 \pm 10.31$ , and  $z = 34.86 \pm 4.89$  for the mPFC, with two participants missing data; in experiment 2 corresponded to:  $x = -50.50 \pm 2.38$ ,  $y = 8.57 \pm 1.34$ , and  $z = 23.29 \pm 0.61$  for vPM, with one missing participant data; and across all experiments  $x = 0.44 \pm 1.357$ ,  $y = -16.55 \pm 1.54$ , and  $z = 65.23 \pm 3.60$  (Talairach coordinates) with two missing data for Vertex, which was selected as a control site for the stimulated areas (Jung et al., 2016). This experimental design ensured that any observed effects were specific to the mPFC stimulation rather than a general result of TMS application. Indeed, the Experiment 2 itself served as a control experiment, with Vertex consistently used as the control site, while the left vPM was the region of interest. Importantly, no modulation of bias was expected from stimulation over vPM, making it an ideal control region to compare against the mPFC effects.

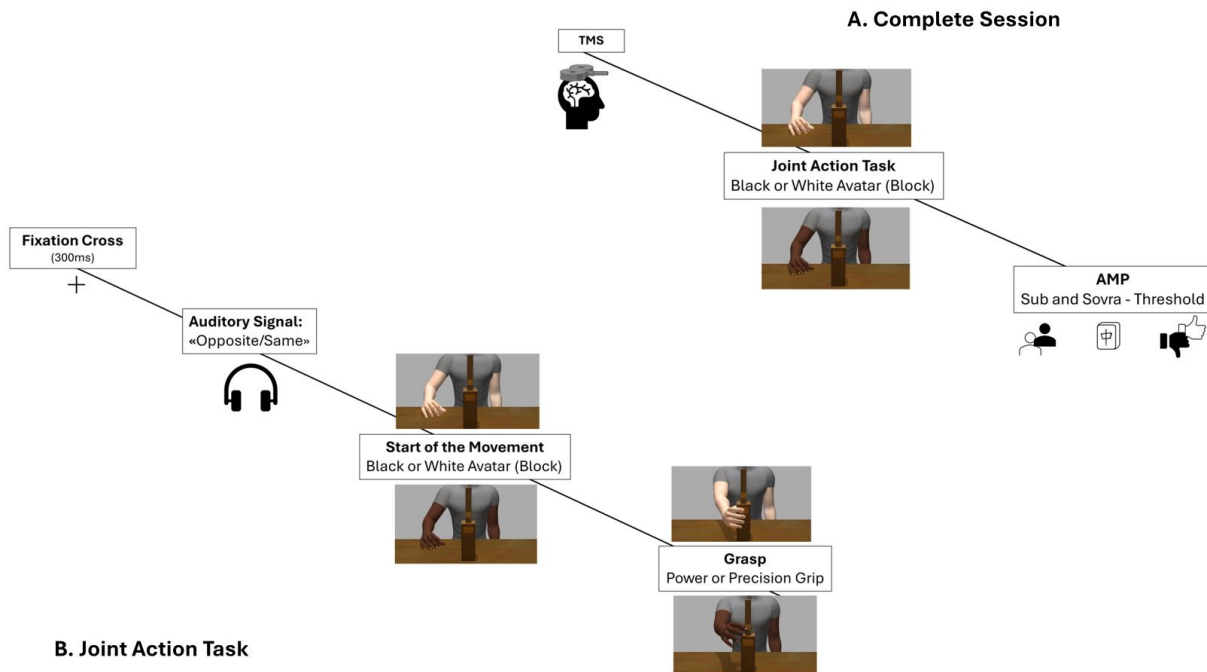
### **Cover story**

To induce participants to perceive the motor interaction with the two virtual partners as more realistic and underline their ethnic differences, a cover story was presented to them at the beginning of the experimental session. Participants were told they would be involved in a study about motor

coordination, in which they had to synchronize with a partner to grasp a bottle-shaped object. They were presented with some tutorial videos showing two individuals performing the task and were told that the experiment had been previously conducted on dyads composed of members of different social groups (e.g., friends, strangers, teammates, and people coming from opposing teams). Notably, it was underlined that in one version of the study, people from the same or different ethnic groups were involved in the coordination task. Crucially, participants were made aware that they were about to interact with two virtual partners whose movements reproduced the kinematics recorded during the task from two real participants. One partner was described as a white-skinned Italian guy named Luca, born in Rome 24 years old at the time of the experiment (in-group). The other partner was presented as a black-skinned Nigerian guy named Ibrahim, who, at the time of the experiment, was 25 years old and lived in Rome for five months (out-group).

## **Procedure**

The experimental procedure was similar among the three experiments, with the only exception being the stimulation sites and the stimulation threshold. At first, participants arrived in the lab, signed the informed consent, and filled out the screening questionnaire for TMS and the handedness inventory. Following that, the experimenters exposed the cover story to the participants while instructing them about the rules of the interactive task. After the explanation, participants completed the two versions of the AMP for the first time. The Sub-threshold AMP version was completed before the Sovra-threshold one. Each version included 72 trials. At this point, the procedure of neuronavigation began, and the first brain stimulation session took place. After the stimulation, participants were required to wait without speaking or standing for 4 minutes. Subsequently, the joint grasping task started, and participants interacted with each virtual partner (Black/White) in a separate block composed of 64 trials. The order of the blocks was counterbalanced between participants. When the second block was completed, participants completed again the two versions of the AMP in the same order (Sub-threshold preceding Sovra-threshold) and were asked to wait in the laboratory for some time until 1 hour from the previous stimulation session was passed. Then, the neuronavigation was conducted again, and TMS was applied over the remaining stimulation site (the order of the stimulation site was balanced between participants, even in this case). Participants underwent the two blocks of joint grasping and finally completed the last two blocks of AMP. Eventually, participants were debriefed about the experimental hypotheses and cover story.



**Fig. 3. A) Complete Session:** The experimental session began with the TMS stimulation, followed by the Joint Action Task in which participants interacted with either the Black or White avatar across blocks in a counterbalanced order. Afterward, participants completed both Sub- and Sovra-threshold versions of the AMP task. **B) Joint Action Task:** In each trial, participants first saw a fixation cross (300 ms), followed by an auditory signal indicating whether they had to perform a "same" or "opposite" action relative to the avatar. Participants then executed the grasping movement with either the Black or White avatar depending on the block.

## Statistical Analyses

Before performing the statistical analyses related to the joint grasping task, the raw data were cleaned as follows: 1) trials in which participants failed to grasp the bottle-shaped object correctly were excluded from the analyses (i.e., when the touch-sensitive plates were not correctly pressed), 2) when participants did not follow the instructions, and grasped the object on the wrong half the trial was discarded, 3) in a similar way, we did not include in the analyses those trials in which participants started their movement before the audio instruction was given, 4) finally we excluded catch trials (i.e., when the virtual partner switched movement direction) from the analyses.

For the joint grasping task, the independent variables included:

- Site (mPFC vs vPM vs Vertex)
- Virtual partner Ethnicity (Black vs White)
- Movement (Same vs Opposite)
- Experiment (1, 2 or 3)

The dependent variables taken into account for the Joint Action task are the following ones:

- Grasping Asynchrony (GA), i.e., the absolute value of the difference between the time of the thumb-index contact on the bottle-shaped object for the virtual partner and the participant.
- Movement Time (MT), i.e., the difference between the time the participant released the starting button and the time of their thumb-index contact with the object.
- Reaction Time (RT), i.e., the time passed between the end of the audio instruction and the moment the participant released the starting button.

To avoid extreme values obtained in some trials that could bias the estimation of the effects of interest, we calculated the mean reaction time (RT) for each subject per Site, Ethnicity, and Movement. For the other two dependent variables (GA, MT), we also included the factor of Grasp Type (Precision/Power). Outliers were then excluded, defined as trial values 2.5 standard deviations above or below the mean for each subject within each condition.

Additional analyses were performed on the AMP (separately for the Sovra and Sub-threshold versions). In this case, the independent variables were:

- Site (mPFC vs vPM vs Vertex)
- Ethnicity of the prime (Black vs White)
- Session (1 (i.e., baseline), 2 or 3)
- Experiment (1, 2 or 3)

In this instance, the dependent variable was represented by the score reflecting the evaluation of the ideograms presented during the AMP. To compute the score, each segment of the VAS was associated with a specific value from 0 to 100. 0 corresponded to the last segment at the bottom end of the VAS and indicated that the participant evaluated the ideogram as “not at all” pleasant. Conversely, 100 corresponded to the last segment at the top end of the VAS, indicating that the participant evaluated the ideogram as “extremely” pleasant.

As for data from the Joint Action task, outlier values were excluded before performing statistical analyses. The mean for both scores was calculated per Site, Ethnicity, Session, and Experiment for each subject. Scores of 2.5 standard deviations below or above the mean for each subject within each condition and level of the independent variables were considered outliers and were subsequently discarded (0.43% of the total number of scores for Sovra-threshold AMP, 0.47% of the total for Sub-threshold AMP).

A series of Shapiro-Wilk normality tests conducted on the R software (Kassambara 2020, package rstatix) indicated that the distribution of all dependent variables from the AMP and the Joint

Grasping Task diverged from normality. Due to these deviations, the application of standard ANOVA was not appropriate for our data and even non-parametric approaches, such as the Aligned Rank Transform (ART) used in the ARTool package (Wobbrock et al., 2011), presented challenges given the complexity and distributional characteristics of our dataset. Moreover, we wanted to account for intersubject variability in their bias toward ethnic in/out-groups and variability in the effects exerted by TMS on participants' performance. Therefore, data from the AMP and the Joint Grasping tasks were analyzed using Multilevel Linear Mixed Models (MLMM) with the R software (Development Core Team 2013, package 'lme4').

In constructing the MLMMs, we performed a "stepwise backward" selection process for the random effects. Starting from a structure that included all the relevant effects and interactions, we systematically removed terms based on singularity to optimize model fit and ensure the convergence of the model. Post-hoc comparisons were conducted utilizing the 'Estimated Marginal Means' R package (Lenth, 2020), employing the 'emmeans' and 'emtrends' functions to further investigate significant interactions. Bonferroni correction for multiple comparisons was applied.

### **2.4.3. Results**

#### **AMP Results**

In analysing Affect Misattribution Procedure (AMP) data we employed linear mixed-effects models for both Sub-threshold and Sovra-threshold AMP. First, to investigate the presence of a racial bias at baseline we combined data from the three experiments, kept only the trials from the baseline sessions and built two models (one for Sub- and one for Sovra-threshold data). Those models included Visual Analog Scale (VAS) scores as the dependent variable, while the main effects of Experiment and Ethnicity, along with their interaction represented the fixed effects. Due to singularity issues, the random effects structure was simplified to include only the random intercept by subject across both models.

Additional models were built to estimate the effects of the TMS on VAS scores. In this case, it was not possible to analyse data from the three experiments merged, due to a different site of stimulation in experiment 2 compared to experiments 1 and 3. Indeed, applying a single model to a dataset including all 3 experiments would have resulted in rank deficiency, since the "vPM" level of the Site factor and the "2" level of the Experiment factor were always aligned. Consequently, we decided to implement two models for both Sovra- and Sub-threshold data, one including

experiments 1 and 3, and one including only experiment 2. As the dependent variable, in those models, we included an index computed for each subject by subtracting the mean VAS score for each stimulation site (i.e., Vertex, mPFC or vPM) from the VAS score at baseline (e.g., mean VAS baseline - mean VAS mPFC). In this way, positive values of the index entailed that a given stimulation site produced a negative bias compared to baseline, while negative values of the index meant that a given stimulation site induced a positive bias in the evaluation. The fixed effects structure for the models including study 1 and 3 comprehended: the main effects of Session, Experiment and Site, the interaction between Site and Ethnicity, the one between Site and Experiment, and the triple interaction between Site, Ethnicity and Experiment. The random effect structure, due to singularity, included only the random effect of Site and the random intercept by subject.

For the models related to experiment 2, the fixed effect structure was composed of the main effects for Site and Session, plus the interaction between Site and Ethnicity. Again, to avoid singularity, the random effects structure was simplified to include only the random intercept by subject.

The residuals of the linear mixed models were assessed to ensure normality. A visual inspection of the residual density plot indicated an approximately normal distribution. Furthermore, diagnostic plots, including residual vs. fitted values, were examined using the function of 'qqnorm' and 'qqplot' in R. Also these plots supported overall the assumption of normality, as well as independence and homoscedasticity, across all models.

### **Sub-threshold AMP**

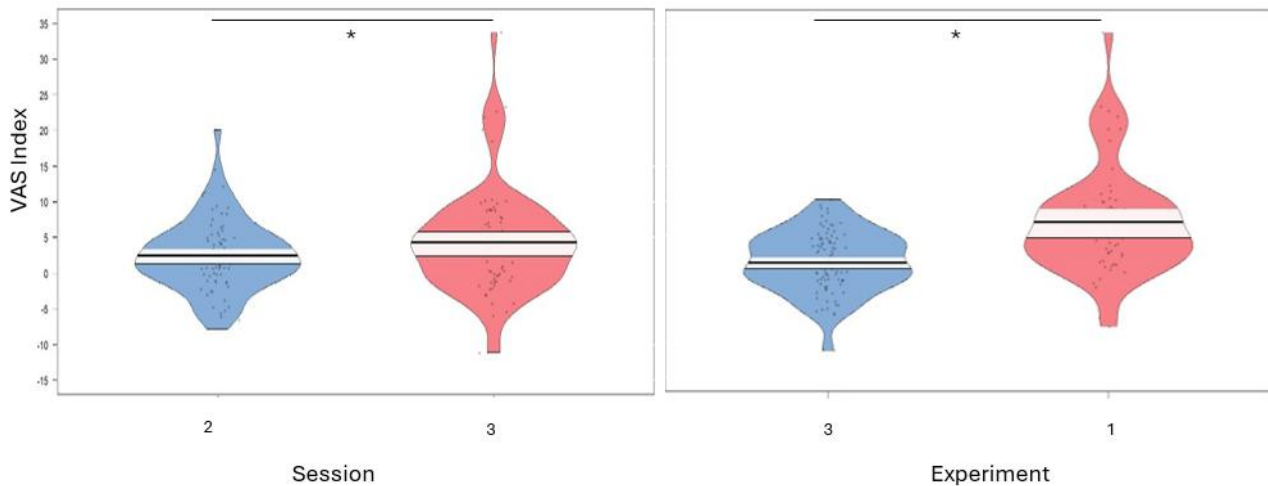
Here, the combined results from Experiment 1 and 3 are presented separately from Experiment 2.

#### *Experiment 1 and 3*

Discussing the results of the Sub-threshold AMP, the output of the model used to estimate the racial bias at baseline indicated that the main effects of Experiment ( $\chi^2 = 4.41$ ,  $p = .11$ ) and Ethnicity ( $\chi^2 = 1.20$ ,  $p = .27$ ), and their interaction ( $\chi^2 = .61$ ,  $p = .74$ ) were not significant.

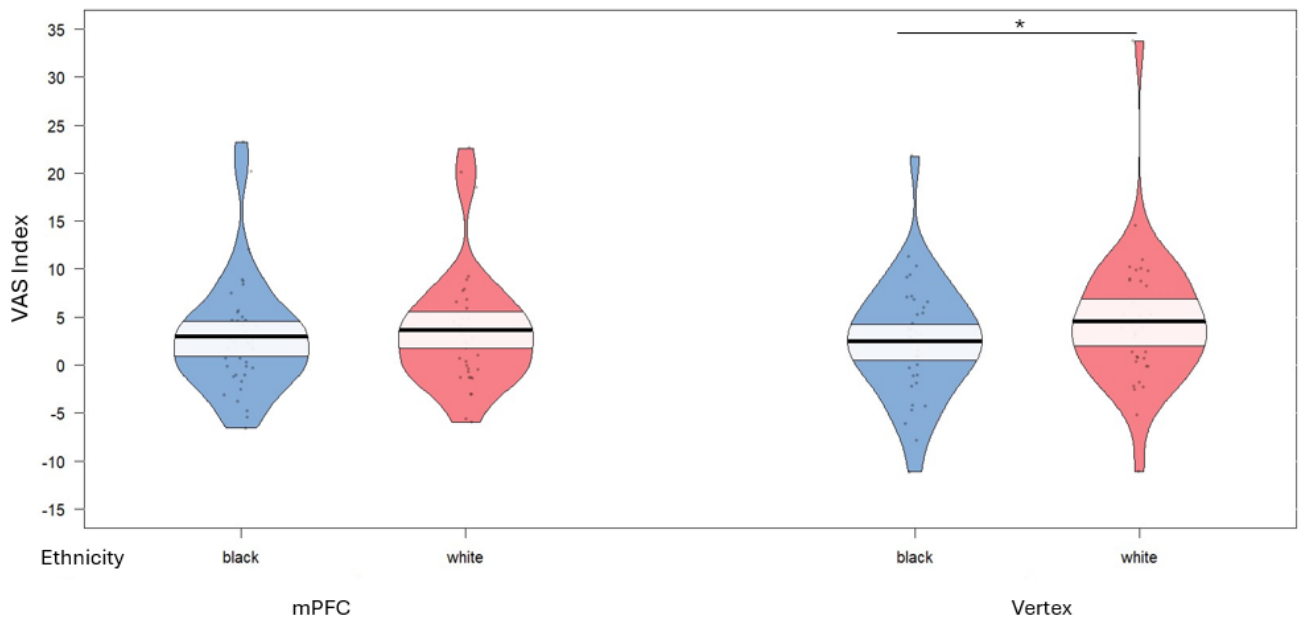
Results related to the model in which the effect of the stimulation site was estimated on the dataset combining Experiments 1 and 3, highlighted the presence of significant main effects for Session ( $\chi^2 = 5.66$ ,  $p = .02$ ) and Experiment ( $\chi^2 = 14.16$ ,  $p < .001$ ). In Particular, the VAS index was higher in session 3 compared to session 2 ( $\beta = 1.93$ ), revealing a more positive evaluation of ideograms in the third session than in the second (Fig. 4, right graph). The index was also higher in Experiment 1

compared to experiment 3 ( $\beta = 5.71$ ), underlying a more positive evaluation in the first experiment than in the third one (Fig. 4, left graph).



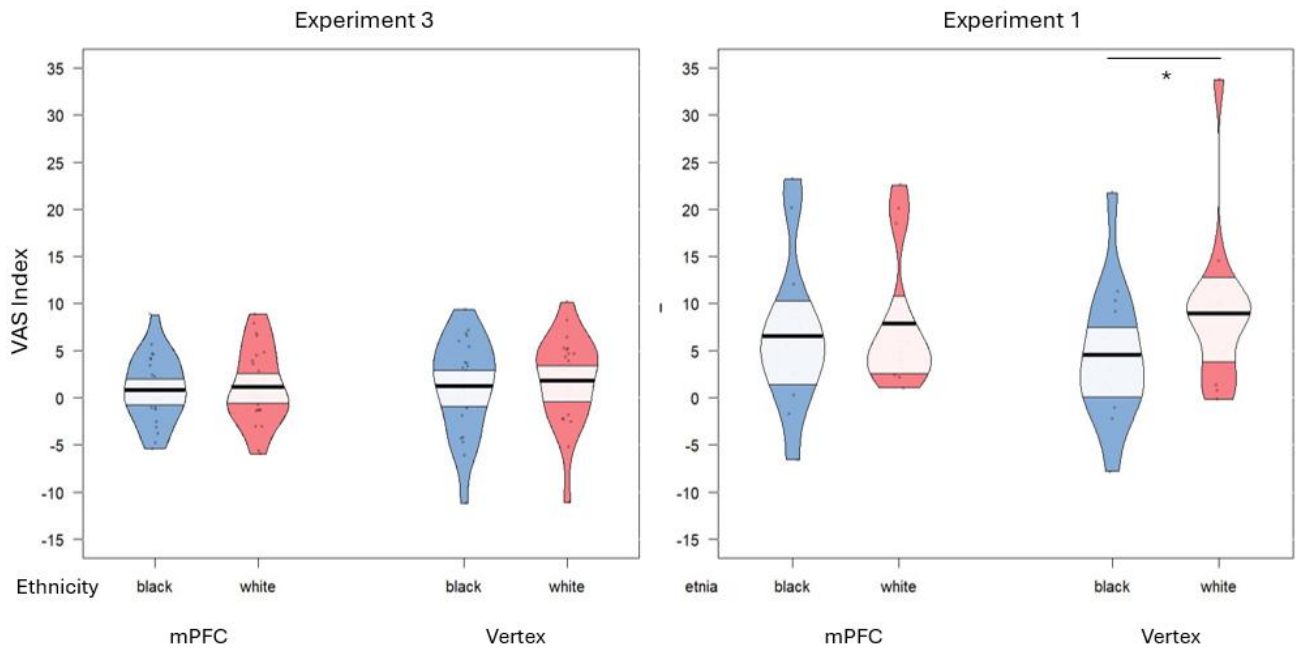
**Figure 4. Results from the Sub-threshold AMP LMM on the VAS index for racial bias.** The graph on the right presents the significant increase in the VAS index during Session 3 compared to Session 2, indicating a more positive evaluation of ideograms in the third session. The left graph shows a notable difference between Experiments 1 compared to 3, suggesting that participants rated the ideograms more favorably during the first experiment compared to the third.

Moreover a significant interaction between Ethnicity and Site ( $\chi^2 = 10.90$ ,  $p < .01$ ) emerged. Post-hoc analyses showed that in the Vertex stimulation condition the index was higher when ideograms were preceded by white compared to black faces ( $\beta = 2.51$ ,  $t = 3.74$ ,  $p < .01$ ), while this effect was not present in mPFC stimulation condition ( $\beta = .80$ ,  $t = 1.20$ ,  $p < .63$ ) (Fig. 5). This entails that vertex stimulation, compared to baseline, produced a stronger increase in evaluation of ideograms preceded by white faces than those preceded by black (i.e., increased pro-white bias). Differently, when mPFC was stimulated, the change from baseline in judgment of ideograms preceded by white and black faces was similar (i.e., no change in racial bias).



**Figure 5. Interaction effects of Ethnicity and Stimulation Site of AMP Subthreshold on the VAS index for racial bias.** Vertex stimulation condition (on the right), showed a substantial increase in the VAS index when ideograms were preceded by white faces compared to black. This difference was not found for mPFC.

Finally, a significant triple interaction of Ethnicity, Site and Experiment ( $\chi^2 = 8.78$ ,  $p = .01$ ) was found. Post-hoc comparisons revealed that in the first experiment, but not in the third, in the vertex condition the index was higher when ideograms were preceded by white compared to black faces ( $\beta = 4.44$ ,  $t = 4.18$ ,  $p < .001$ ) (Fig. 6). Thus, the previously reported significant interaction between Site and Ethnicity emerged only in the first experiment. All other comparisons were not significant (all  $ps > .05$ ). The main effect of Site ( $\chi^2 = .02$ ,  $p < .88$ ) and the interaction between Experiment and Site ( $\chi^2 = .31$ ,  $p = .58$ ) were not significant.



**Figure 6. Interaction effects of Experiment, Ethnicity and Stimulation Site of AMP Subthreshold on the VAS index for racial bias.** Only in Experiment 1 (on the right) Vertex stimulation condition showed a substantial increase in the VAS index when ideograms were preceded by white faces compared to black. This difference was not found for mPFC of both Experiments or Vertex in Experiment 3.

### *Experiment 2*

The last model related to the Sub-threshold AMP was performed on data from experiment 2. Results indicated no significant effects of Site ( $\chi^2 = .74$ ,  $p = .39$ ), Session ( $\chi^2 = .73$ ,  $p = .39$ ) and of the interaction between Site and Ethnicity ( $\chi^2 = .09$ ,  $p = .95$ ).

### **Sovra-threshold AMP**

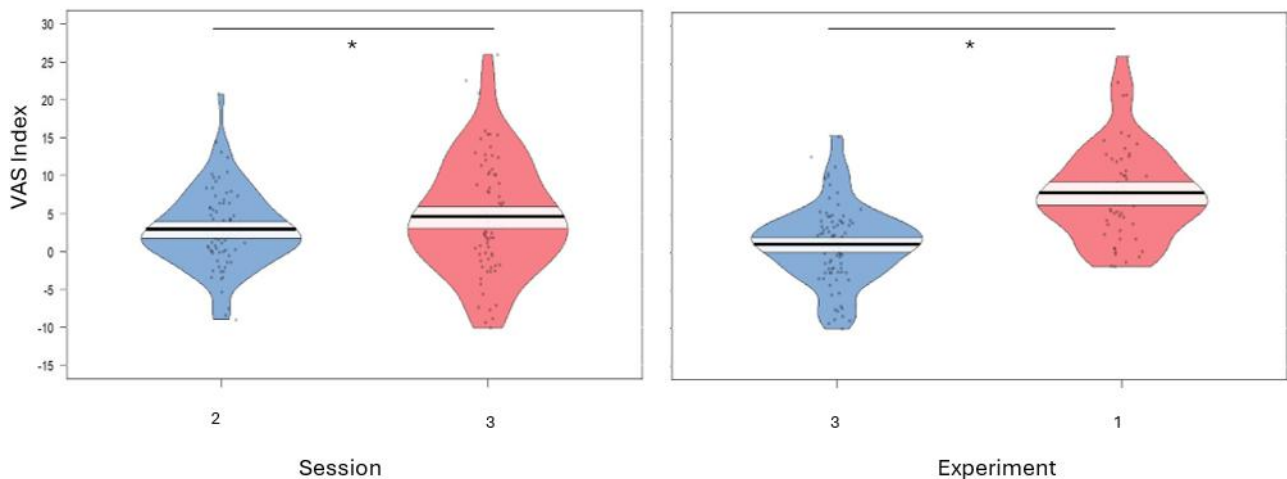
Here, the combined results from Experiment 1 and 3 are presented separately from Experiment 2.

### *Experiment 1 and 3*

Moving to the Sovra-threshold AMP, results from the first model, which included data for the baseline session from the three experiments, showed that the main effects of Experiment ( $\chi^2 = 1.87$ ,  $p = .39$ ) and Ethnicity ( $\chi^2 = 1.89$ ,  $p = .17$ ) were not significant, and neither was the interaction between Ethnicity and Experiment ( $\chi^2 = 1.79$ ,  $p = .40$ ).

Concerning the model used to estimate the effects of the stimulation site on the dataset including experiments 1 and 3, significant main effects of Session ( $\chi^2 = 5.46$ ,  $p = .02$ ) and Experiment ( $\chi^2 = 19.32$ ,  $p < .001$ ) emerged (Fig. 7, graph on the right). Specifically, the index was higher in Session 3 than in Session 2 ( $\beta = 1.64$ ), indicating a more positive evaluation of ideograms in the third

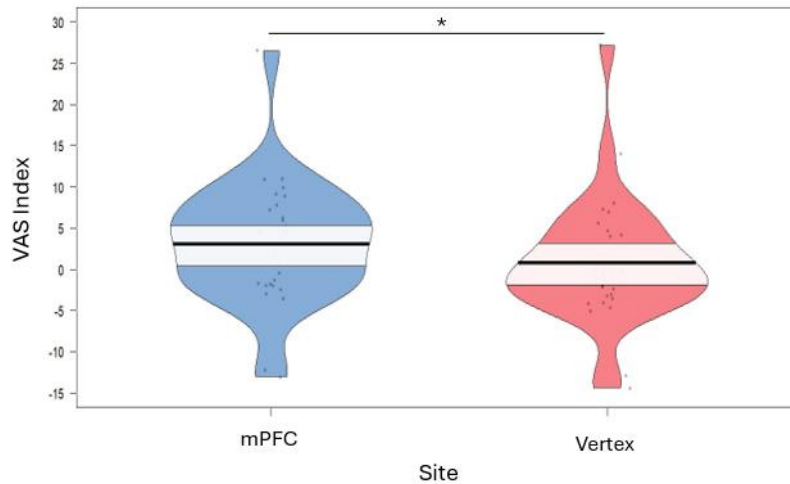
session compared to the second. In addition, the index was higher in Experiment 1 compared to 3 ( $\beta = 6.46$ ), entailing a more positive evaluation in the first experiment than in the third one (Fig. 7, graph on the left). Overall these findings complement the one found for AMP Sub-threshold. The main effect of Site ( $\chi^2 = .01$ ,  $p = .93$ ) and all of the interaction effects (all  $p$ s  $> .50$ ) did not overpass the significance threshold.



**Figure 7. Results from the Sovrathreshold AMP LMM on the VAS index for racial bias.** The graph on the right presents the significant increase in the VAS index during Session 3 compared to Session 2, indicating a more positive evaluation of ideograms in the third session. The graph on the left shows a notable difference between Experiments 1 compared to 3, suggesting that participants rated the ideograms more favorably during the first experiment compared to the third.

### Experiment 2

The last computed model for the Sovra-threshold AMP was performed on data from the second experiment, again to estimate the effects related to the stimulation site. Results revealed a significant effect of Site ( $\chi^2 = 4.95$ ,  $p = .03$ ) (Fig. 8). The VAS index was higher for vertex than vPM stimulation ( $\beta = 2.50$ ), underlying a more positive evaluation of ideograms after TMS over vertex compared to TMS over vPM. The main effect of Session ( $\chi^2 = 3.05$ ,  $p = .08$ ) and the interaction between Site and Ethnicity ( $\chi^2 = 0.46$ ,  $p = .80$ ) were not significant.



**Figure 8. Results from the Sovrathreshold AMP LMM on the VAS index for racial bias.** The graph illustrates the significant main effect of Site on the VAS index showing a higher positive evaluation after mPFC stimulation for the ideograms.

### Joint Action Task Results

In the following sections, we report the results from the MLMM on the three experiments in the joint action task for Grasping Asynchrony, Movement Time, and Reaction Times. We initially conducted analyses on the combined data from experiments 1 and 3, and separately on data from experiment 2. Given the small sample sizes, however, we then decided to examine the impact of our factors in the joint action task on the dependent variables across the entire dataset. In these latter analyses across all studies, we included 'Site' as a covariate for GA, since a Site-specific effect was observed in the individual studies for this dependent variable. Conversely, 'Site' was not included in the three studies combined analyses for MT and RT, as no site-specific effects were found in the individual studies for these variables.

As for AMP, the residuals of the linear mixed models were assessed to ensure normality. A visual inspection of the residual density plot indicated an approximately normal distribution. Furthermore, diagnostic plots, including residual vs. fitted values, were examined using the function of 'qqnorm' and 'qqplot' in R. Also these plots supported overall the assumption of normality, as well as independence and homoscedasticity, across all models

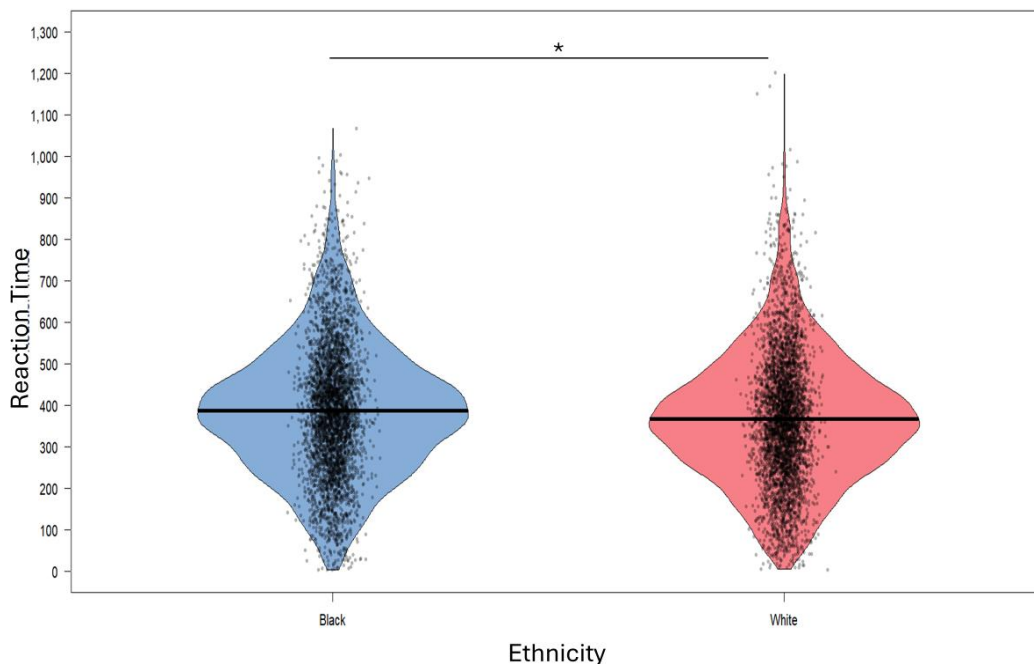
## Experiment 1 and 3

Here, the combined results from Experiment 1 and 3 are presented.

### *Reaction Times*

The analyses involved the factors of Experiment (1, 3) \* Ethnicity (Black, White) \* Site (mPFC, V), Movement (Same, Opposite) on Reaction Times, considering together with the the random intercept by subject, the interaction of Ethnicity \* Site in the random structure. A main effect of Ethnicity ( $\chi^2 = 7.38$ ,  $p = 0.01$ ) was found, suggesting that the time between participants and the avatar's grasp of the object was different depending on the group membership (Fig. 9). The interaction between Experiment vs. Movement vs. Ethnicity was significant ( $\chi^2 = 4.23$ ,  $p = 0.04$ ). No other main effect or interaction was found significant (all  $ps > 0.05$ ).

At the posthoc level, the main effect of Ethnicity revealed that the grasping movement led to significantly slower RTs for Black ( $M = 386.92$ ,  $SD = 60.27$ ) compared to the interactions with the White avatar ( $M = 366.93$ ,  $SD = 159.05$ ). In contrast, all the other comparisons resulted not significant (all  $ps > 0.05$ ).



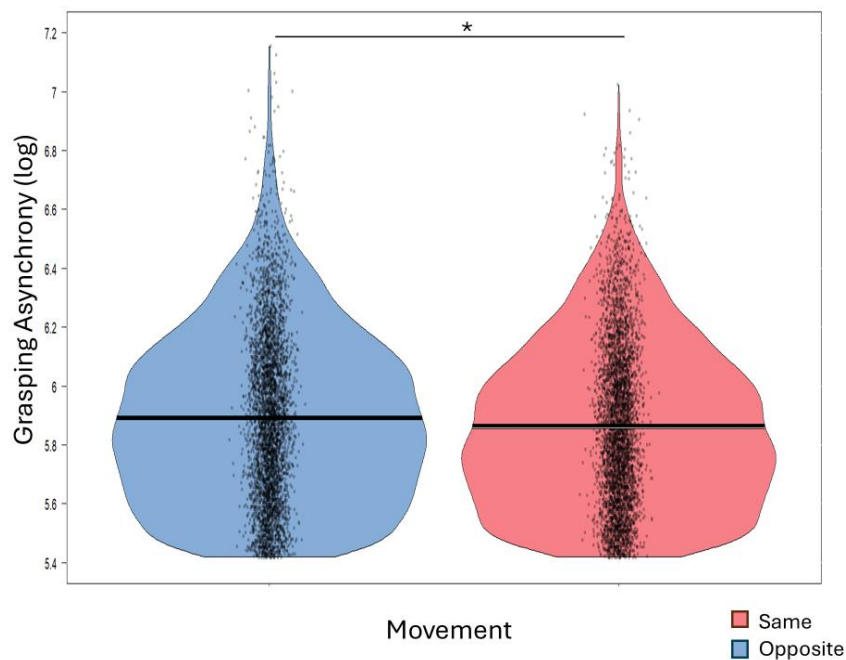
**Figure 9. Results from the Joint Action in Experiment 1-3 from LMM on the RTs.** The significant main effect of Ethnicity indicates differences in reaction times based on group membership, with RTs significantly slower for the outgroup interactions.

### *Grasping Asynchrony*

To address violations of the normality assumption, we applied a logarithmic transformation to the GA values. This transformation allowed to better approximate normality, making the data more suitable to the LMM analysis. Following the transformation, the same analysis was conducted on the log-transformed GA values.

The model involved the factors of Experiment (1, 3) \* Ethnicity (Black, White) \* Site (mPFC, V), Movement (Same, Opposite) on Grasping Asynchrony, considering the interaction of Ethnicity \* Site in the random structure together with the the random intercept by subject. It revealed a significant main effect of Movement ( $\chi^2 = 22.26$ ,  $p < 0.00$ ) and a significant interaction of Experiment per Site ( $\chi^2 = 4.91$ ,  $p = 0.03$ ) (Fig. 10). No other main effect or interaction was found significant (all  $p > 0.05$ ).

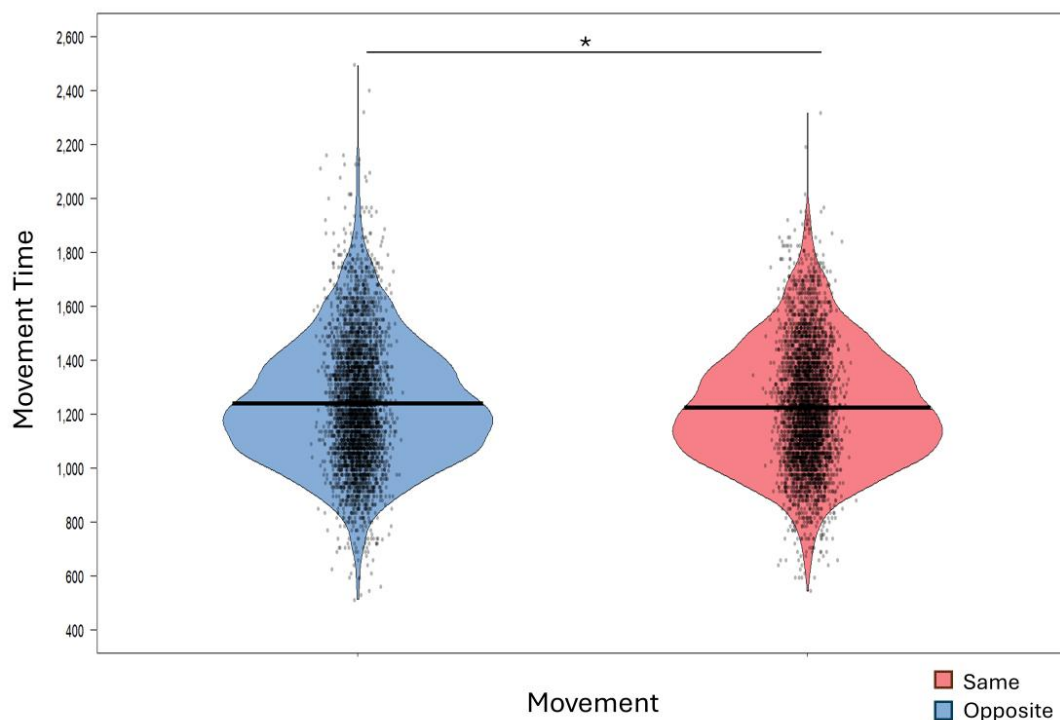
Post-hoc revealed no significant difference emerged in the comparison between the factor Experiment vs. Site levels. Investigating the main effect of Movement, it is shown that the execution of Opposite movement led to a significantly higher Grasping Asynchrony (M = 5.89, SD = 0.29) compared to the Same ones (M = 5.86, SD = 0.28).



**Figure 10. Results from the Joint Action task in Experiment 1-3 from LMM on the GA.** The significant main effect of Movements indicates differences in grasping asynchrony based on the movement performed relative to the partner, with performance significantly worse for Opposite conditions.

### *Movement Time*

The MLMM involved the factors of Experiment (1, 3) \* Ethnicity (Black, White) \* Site (mPFC, V), Movement (Same, Opposite) on Movement Time, considering the interaction of Ethnicity \* Site and the random intercept of subjects in the random structure. The analysis highlighted only a main effect of Movement ( $\chi^2 = 8.27$ ,  $p = 0.04$ ), suggesting that the time between the movement start and grasp of the object differed depending on the executed grasping movement (Fig. 11). In contrast, all other effects and interactions are not significant (all  $p$ s > 0.05). The execution of the Same Movements led to a significantly faster time ( $M = 1225.14$ ,  $SD = 222.60$ ) than the Opposite ones ( $M = 1238.265$ ,  $SD = 233.52$ ).



**Figure 11. Results from the Joint Action task in Experiment 1-3 from LMM on the MT.** The significant main effect of Movements indicates differences in the time of the movement execute were based on the movement itself being performed relative to the partner, with performance significantly slower for Opposite conditions

## Experiment 2

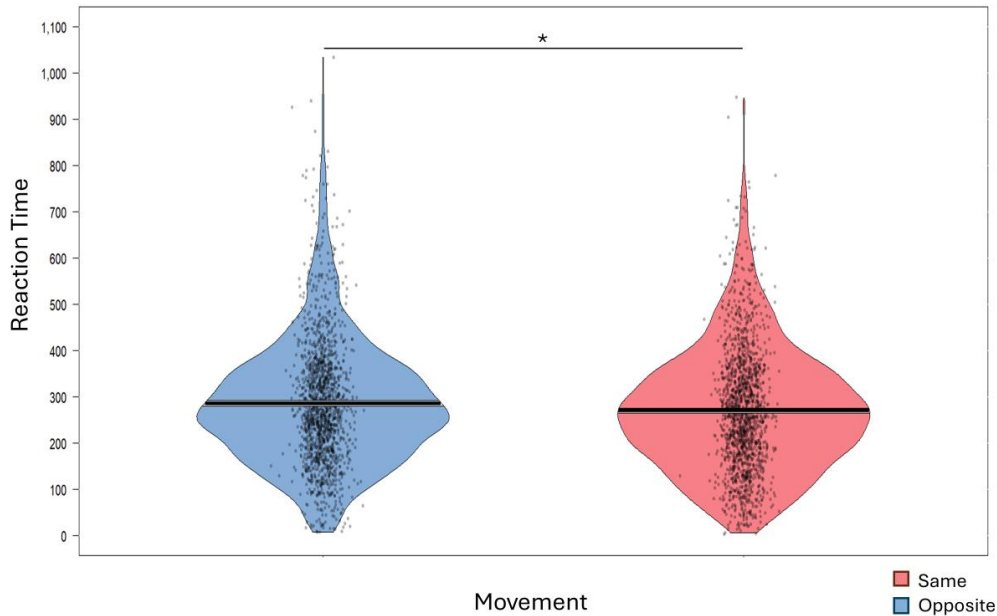
Here are reported the results from Study 2.

### *Reaction Times*

The model involved the factors of Ethnicity (Black, White) \* Site (mPFC, V), Movement (Same, Opposite) on Reaction Times, considering the random intercept by subject and the interaction of Ethnicity \* Site in the random structure. A main effect of Movement ( $\chi^2 = 22.57$ ,  $p < 0.00$ ) was

found, indicating that the time between participants and the avatar's grasp of the object was different depending on the grasping action (Fig. 12). The other results were all not significant (all  $p > 0.05$ ).

The main effect of Movement revealed that the grasping led to significantly slower RT for Opposite movements ( $M = 286.13$ ,  $SD = 138.08$ ) compared to the Same ones ( $M = 270.78$ ,  $SD = 135.47$ ).

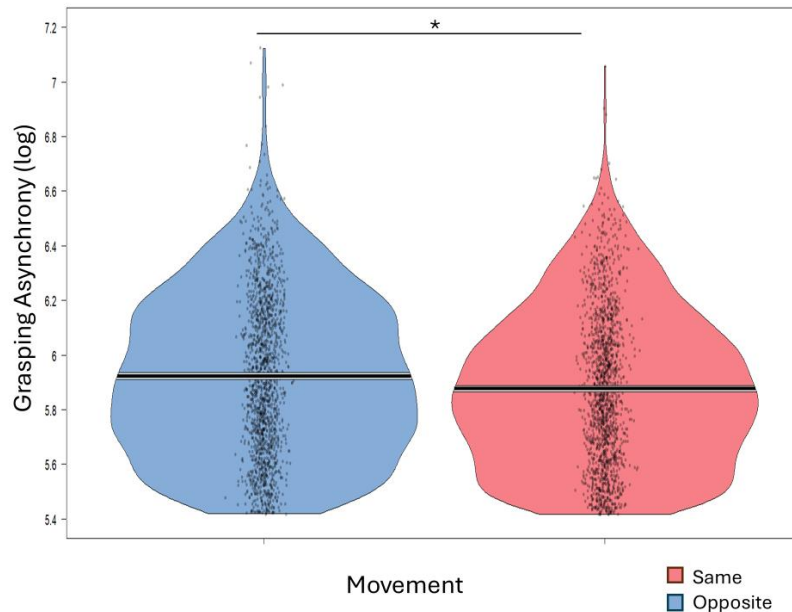


**Figure 12. Results from the Joint Action in Experiment 2 from LMM on the RTs.** The significant main effect of Movement indicates differences in reaction times based on the action performed relative to the partner, with RTs significantly slower for the Opposite interactions.

### *Grasping Asynchrony*

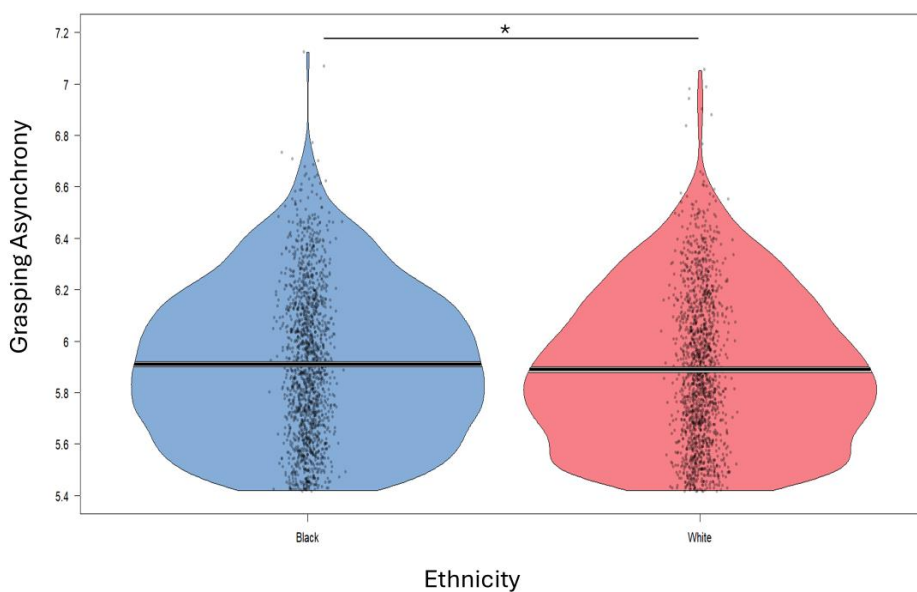
To address violations of the normality assumption, we applied a logarithmic transformation to the GA values. This transformation allowed to better approximate normality, making the data more suitable to the LMM analysis. Following the transformation, the same analysis was conducted on the log-transformed GA values.

The model involved the factors of Ethnicity (Black, White) \* Site (vPM, V) \* Movement (Same, Opposite) on Grasping Asynchrony, considering only the random intercept by subject in the random structure. It revealed a significant main effect of Ethnicity ( $\chi^2 = 4.47$ ,  $p = 0.03$ ) and a main effect of Movement ( $\chi^2 = 22.54$ ,  $p < 0.00$ ) (Fig. 13). No other main effect or interaction was found significant (all  $p > 0.05$ ). The main effect of Movement revealed that the grasping led to significantly greater asynchrony for Opposite ( $M = 5.92$ ,  $SD = 0.30$ ) compared to the Same movements ( $M = 5.88$ ,  $SD = 0.28$ ).



**Figure 13. Results from the Joint Action in Experiment 2 from LMM on the GA.** The significant main effect of Movement indicates differences in grasping asynchrony based on the action performed relative to the partner, with performance significantly worse for the Opposite interactions.

The main effect of Ethnicity revealed that the grasping led to significantly greater asynchrony when interacting with the Black avatar ( $M = 5.91$ ,  $SD = 0.29$ ) compared to the White one ( $M = 5.89$ ,  $SD = 0.29$ ).



**Figure 12. Results from the Joint Action in Experiment 2 from LMM on the GA.** The significant main effect of Ethnicity indicates differences in grasping asynchrony based on the partner group membership, with performance significantly worse for the outgroup interactions.

### *Movement Time*

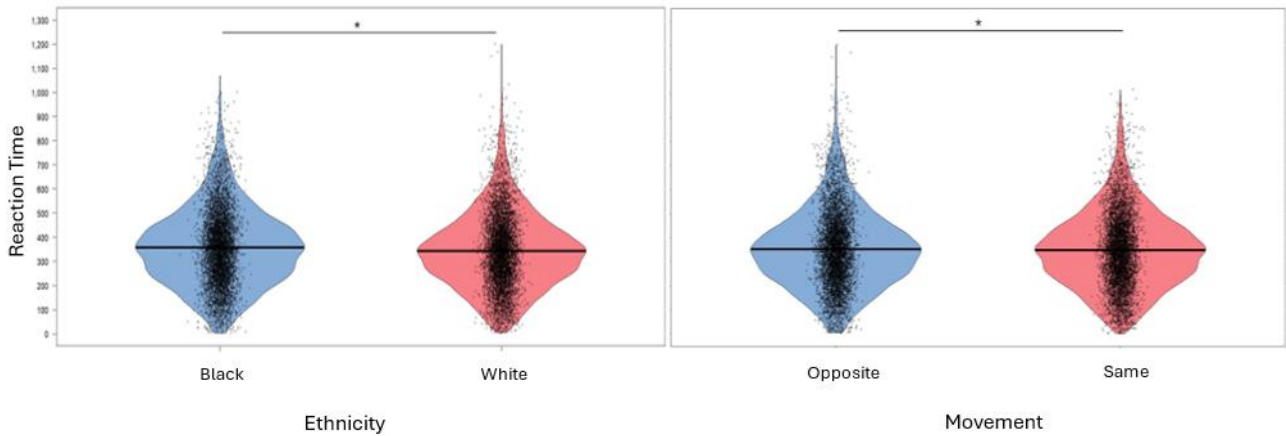
The analysis involved the factors of Ethnicity (Black, White) \* Site (mPFC, V) \* Movement (Same, Opposite) on Movement Time, considering the interaction of Ethnicity \* Site in the random structure together with the the random intercept by subject. The results were all not significant (all  $p$ s > 0.05).

### **All Experiments**

Here are presented the combined results from Experiment 1 to 3.

### *Reaction Times*

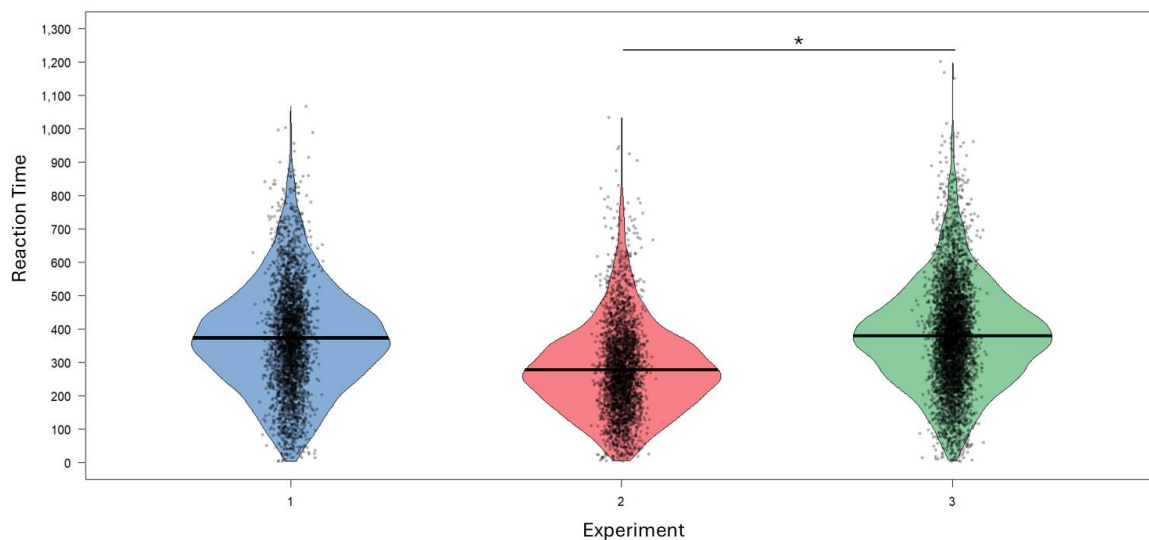
The analysis involved the factors of Experiment (1, 2, 3) \* Ethnicity (Black, White) \* Movement (Same, Opposite) on Reaction Times, considering the random intercept by subject and the factor of Ethnicity in the random structure. The results highlighted the main effect of all three factors, indicating that the time between participants and the avatar's grasp of the object was different depending on the Experiment performed ( $\chi^2 = 12.95$ ,  $p = 0.00$ ), the avatar Ethnicity ( $\chi^2 = 7.21$ ,  $p = 0.01$ ) and the executed grasping Movement ( $\chi^2 = 3.89$ ,  $p = 0.05$ ). Furthermore, also the interaction Experiment \* Movement resulted significant ( $\chi^2 = 13.03$ ,  $p = 0.00$ ). All other effects and interactions were not significant (all  $p$ s > 0.05). The execution of the task with the White avatar led to a significantly faster time ( $M = 342.91$ ,  $SD = 158.18$ ) compared to the Black one ( $M = 357.21$ ,  $SD = 161.80$ ) (Fig. 13, on the right). The execution of the Opposite Movements led to a significantly slower time ( $M = 352.07$ ,  $SD = 161.32$ ) compared to the Same ones ( $M = 347.99$ ,  $SD = 158.96$ ) (Fig. 13, on the left).

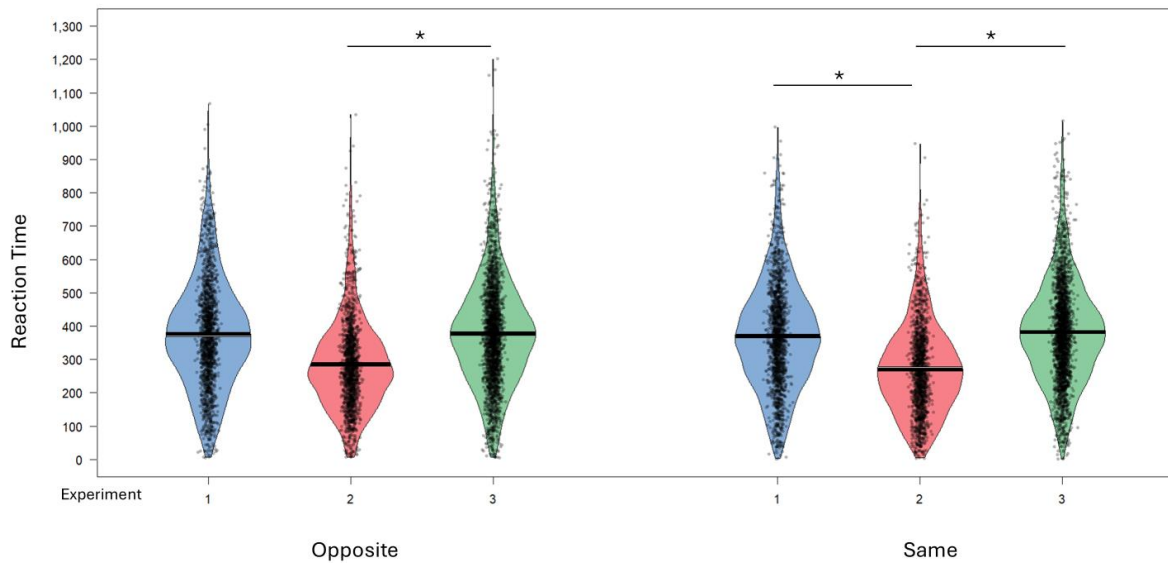


**Figure 13. Results from the Joint Action in all Experiments from LMM on the RTs.** The significant main effect of Ethnicity (on the right) indicates differences in reaction based on the partner group membership, with performance significantly slower for the outgroup interactions. The significant Main effect of Movement (on the left), indicate that RTs are also significantly slower for Opposite movements.

Post hoc revealed that reaction time was significantly faster in Experiment 2 ( $M = 278.40$ ,  $SD = 136.96$ ;  $p = 0.05$ ) compared to the third one ( $M = 379.60$ ,  $SD = 158.13$ ), but neither was significantly different from the first one ( $M = 372.65$ ,  $SD = 162.57$ ) (Fig. 14, upper graph).

RTs were significantly faster for both Opposite ( $p = 0.04$ ) and Same Movements ( $p = 0.05$ ) in Experiment 2 (respectively  $M = 286.13$ ,  $SD = 138.08$ ;  $M = 270.78$ ,  $SD = 135.47$ ) compared to the third one (respectively  $M = 377.62$ ,  $SD = 160.80$ ;  $M = 381.57$ ,  $SD = 155.45$ ). Furthermore, RTs for Same movement were significantly faster in Experiment 2 ( $M = 270.78$ ,  $SD = 135.47$ ) compared to the first one ( $M = 369.96$ ,  $SD = 160.20$ ;  $p = 0.00$ ) (Fig 14, lower graph).





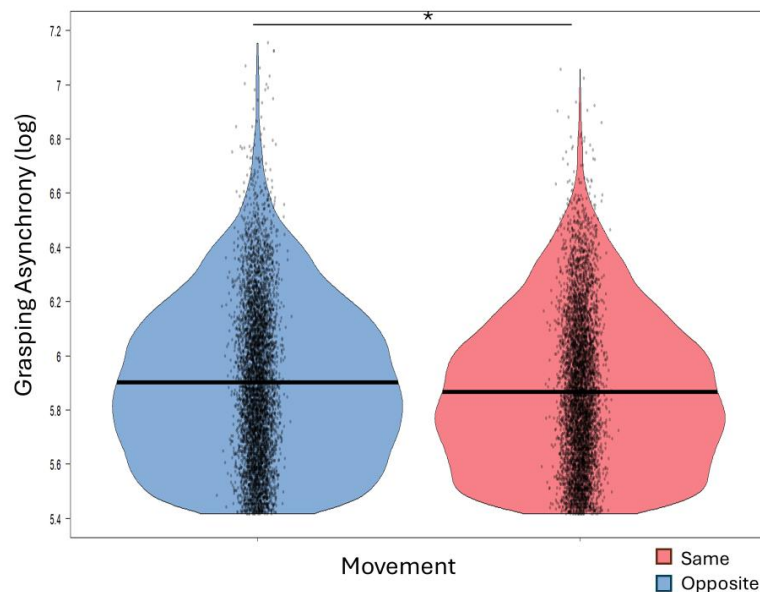
**Figure 14. Results from the Joint Action in all Experiments from LMM on the RTs.** The reaction times were significantly faster in Experiment 2 compared to Experiment 3 (upper graph), while no significant difference was found with Experiment 1. RTs were faster in Experiment 2 compared to Experiment 3 for both Opposite and Same Movements (lower graph). Additionally, RTs for Same Movements were significantly faster in Experiment 2 compared to Experiment 1.

### *Grasping Asynchrony*

To address violations of the normality assumption, we applied a logarithmic transformation to the GA values. This transformation allowed to better approximate normality, making the data more suitable to the LMM analysis. Following the transformation, the same analysis was conducted on the log-transformed GA values.

The model involved the factors of Experiment (1, 2, 3) \* Ethnicity (Black, White) \* Movement (Same, Opposite), with Site (mPFC, vPM, V) as a covariate, on Grasping Asynchrony, considering the factors of Ethnicity and Site in the random structure together with the the random intercept by subject. It revealed a significant main effect of Movement ( $\chi^2 = 42.49$ ,  $p < 0.00$ ) (Fig. 15). No other main effect or interaction was found significant (all  $p > 0.05$ ).

The execution of Opposite movement led to a significantly higher Grasping Asynchrony ( $M = 5.90$ ,  $SD = 0.29$ ) compared to the Same ones ( $M = 5.87$ ,  $SD = 0.28$ ).

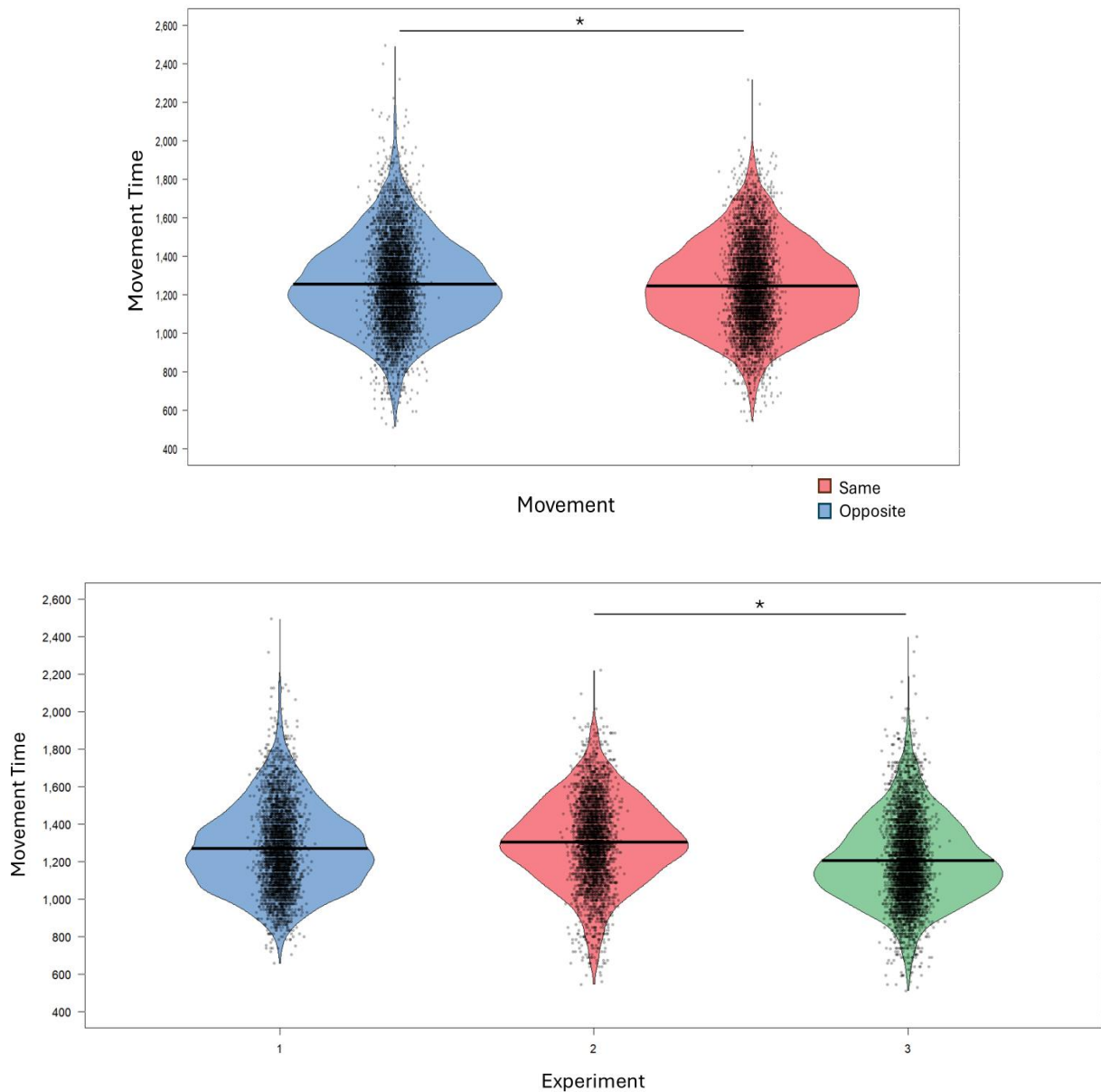


**Figure 15. Results from the Joint Action in all Experiments from LMM on the GA.** The significant main effect of Movement indicates differences in grasping asynchrony with the partner, with performance significantly worse for the Opposite movements.

### *Movement Time*

The analysis involved the factors of Experiment (1, 2, 3) \* Ethnicity (Black, White) \* Movement (Same, Opposite) on Movement Time, considering the factor of Ethnicity in the random structure together with the the random intercept by subject. The results highlighted only a main effect of Experiment ( $\chi^2 = 8.12$ ,  $p = 0.02$ ) and Movement ( $\chi^2 = 5.39$ ,  $p = 0.02$ ), suggesting that the time between the movement start and grasp of the object was different depending on the executed grasping movement and the experiment performed. In contrast, all other effects and interactions were not significant (all  $p > 0.05$ ).

The execution of the Opposite Movements led to a significantly slower time ( $M = 1256.56$ ,  $SD = 235.80$ ) compared to the Same ones ( $M = 1247.07$ ,  $SD = 224.82$ ) (Fig. 16, upper graph). Post hoc revealed that movement time was significantly faster in Experiment 3 ( $M = 1205.01$ ,  $SD = 225.54$ ;  $p = 0.05$ ) compared to the second one ( $M = 1305.37$ ,  $SD = 227.67$ ), but neither was significantly different from the first one ( $M = 1271.33$ ,  $SD = 226.34$ ) (Fig. 16, lower graph).



**Figure 16. Results from the Joint Action in all Experiments from LMM on the MT.** The upper graph shows significantly slower movement times for Opposite Movements compared to Same Movements. The lower graph highlights faster movement times in Experiment 3 compared to Experiment 2, while no significant differences were found between Experiments 1 and 3

#### 2.4.4. Discussion

Humans have a natural tendency to categorize others into social groups, often based on ‘in/out-group’ distinctions. Race is one of the most salient markers for group categorization (Cosmides et al., 2003), influencing cognitive, emotional, and motor processes during social interactions (Gutsell and Inzlicht, 2010; Avenanti et al., 2010) and affecting the coordination of movement between

individuals from different racial groups (Sacheli et al., 2015). In the current study, we aimed to investigate whether the temporary inhibition of mPFC through continuous theta burst TMS could affect motor coordination during joint actions. Specifically, we examined how mPFC inhibition influenced motor behaviors such as grasping asynchrony (GA), movement time (MT), and reaction times (RTs) when interacting with in-group (White) versus out-group (Black) avatars. We further explored whether these effects were linked to the implicit racial biases of participants through the affect misattribution procedure (AMP; Payne et al., 2005), providing insight into how social categorization impacts motor coordination beyond perceptual judgments in interaction. By examining the single and combined results from three experiments, we aimed to better understand the neural and social processes underlying racial biases in joint action dynamics.

## **AMP**

At baseline our participants showed no preference for chinese ideograms primed with white compared to black faces, as measured by the Sub- and Sovra-threshold AMP, providing no support for the presence of a racial bias in our sample, at least at the group level. This result is partially in contrast with the evidence provided by other studies using the AMP to measure racial prejudice, which tend to find a pro-white bias in caucasian participants, although the effect is not always significant (Cooley, Lei & Ellerkamp, 2018; Payne et al., 2005; Steele et al., 2018). However, our finding is coherent with the results of a recent study (Connor et al., 2023), which highlight how implicit evaluations are influenced by multiple biases at the same time, and the effects related to race may be smaller compared to other categorization domains (e.g., gender).

Further analyses conducted on data from the AMP sessions following TMS revealed some effects attributable to the site of stimulation. Particularly, in Experiment 2 there was a significant difference in the evaluation of ideograms depending on which brain region was stimulated. Stimulation over the vertex was associated with a better evaluation of ideograms compared to vPM stimulation. Yet, the effect was found only for the Sovra-threshold AMP and was not dependent on the ethnicity of the prime. Interestingly, an interaction between stimulation site and prime ethnicity emerged for the Sub-threshold task on merged data from Experiment 1 and 3. Specifically, the stimulation of the vertex seemed to enhance the pro-white bias of participants, while no bias was induced by the stimulation of mPFC. Nevertheless, this effect was found only in the Sub-threshold AMP and the triple interaction between Site, Ethnicity and Experiment, indicated it was present only in Experiment 1. Thus, the results do not support our third hypothesis (i.e., mPFC inhibition

increases the negative bias toward the racial out-group). The fact that our participants failed to express a racial prejudice at baseline may explain a similar outcome. Indeed, mPFC is suggested to have a relevant role in the regulation of prejudice expression (Amodio, 2014; Amodio & Frith, 2006; Sellaro et al., 2015), however when people do not show implicit biased attitudes (as in our sample) regulatory mechanisms may become irrelevant. Another explanation could be related to a learning effect during the task. The first session of the AMP was never preceded by TMS, and during this session participants became familiar with the task. This may have helped to be more able to ignore priming stimuli during the sessions that followed brain stimulation, consequently making unreliable the estimation of the racial bias.

At last, significant main effects related to the session and the experiment emerged in both versions of the task, which can be attributable to learning effects and small differences in the composition of the samples in the three experiments.

## **Joint Action Task**

### *Reaction Times*

The analysis of RTs revealed several significant findings. First, participants' were influenced by the Ethnicity of the avatar, with slower RTs when interacting with the Black avatar compared to the White avatar. This finding is consistent with prior research on implicit racial bias (Sacheli et al., 2015a).

Additionally, a significant main effect of Movement type was found, with Opposite movements resulting in slower RTs compared to Same movements, in line with prior work showing that motor coordination between individuals is more efficient when their actions are congruent, as opposed to complementary actions, which require more cognitive and motor adjustment (Kourtis, Sebanz, & Knoblich, 2013; Atmaca, Sebanz, & Knoblich, 2011; Sebanz, Knoblich, & Prinz, 2003, 2005). Notably, we also found a significant interaction between Experiment (1, 2, 3) and Movement type (Opposite vs Same), but also a main effect of Experiment in the combined analysis, indicating that our different experimental conditions might have modulate how efficiently participants coordinate their movements with the avatar.

These findings provide important implications for understanding how social factors, such as racial biases, and motor demands interact to shape joint action performance. The slower RTs observed when interacting with an out-group avatar could reflect a breakdown in predictive motor

simulations, which are crucial for efficient coordination in joint tasks (Sebanz et al., 2006; Sachelti et al., 2015a; Vesper et al., 2010).

### *Grasping Asynchrony*

Across all three experiments, a consistent main effect of Movement type (Same vs. Opposite) was observed on GA. Participants exhibited significantly greater asynchrony when performing Opposite movements compared to Same movements, in line with the RTs results. Interestingly, while we observed an interaction between Experiment and brain Site stimulation, there was no significant effect of ethnicity on GA except for Experiment 2. This suggests that the participants' ability to synchronize their movements with the avatar was uniquely influenced by motor factors (i.e., congruency of movements) and neural stimulation rather than the ethnicity of the avatar. The interaction between Experiment and stimulation Site in Experiment 1 and 3, further suggests that mPFC may play varying roles in processing action coordination during joint tasks. Specifically, the stimulation of mPFC might differentially modulate grasping synchronization in a joint action task, as suggested by existing literature highlighting the role of the mPFC in social cognition and motor control (Amodio et al., 2014, 2006; Amodio & Frith, 2006; Brass et al., 2003; Spengler, von Cramon, & Brass, 2010). However, this effect was not strong and requires further investigation.

### *Movement Time*

Similar to GA and RTs, a significant effect of Movement type (Same vs. Opposite) was found across the analyses on MT, with Opposite movements taking longer to execute than Same movements. A significant main effect of Experiment (1, 2, 3) was also observed, indicating that participants' rapidity in reaching the target differed across experiments, suggesting that subtle changes in task design (i.e., the stimulation site, the difference in the stimulation parameters) or procedural adjustments between the experiments may have influenced participants' performance.

### **Limitations of the study**

Several limitations must be acknowledged in the present study. Firstly, the absence of a racial bias at the attitudinal level in our sample, as measured with AMP, may have hindered the estimation of effects related to the stimulation site in the joint action task. Accordingly, a previous study that reported significant effects of mPFC transcranial stimulation on racial bias (Sellaro et al., 2015), was conducted on a sample where a racial bias was present at baseline.

Secondly, our sample was composed exclusively of Italian Caucasian participants. While this approach controls for confounding variables related to cultural background, in line with previous studies (Contreras-Huerta et al., 2014; Sacheli et al., 2015a; Vala et al., 2012), it limits the generalizability of the findings to other ethnic groups and different populations. It would also be valuable to investigate whether familiarity with the out-group members or longer interaction periods could reduce the observed coordination time. Having these data would have provided a more comprehensive understanding of how racial biases affect joint action at the sensorimotor level.

Finally, the role of the mPFC in modulating these biases remains an open question, and in our specific study one limit regarding this finding could be linked to the sample size in each experiment and the differences in TMS stimulation intensities between the first two experiments and the third one. The present findings on the Experiment factor, seem to suggest that this discrepancy, although based on previous literature and safety considerations (Rosenthal et al., 2009; Yazar et al., 2017; Amoruso et al., 2018; Era et al., 2021), introduces a potential confound when comparing the results across our experiments.

### **Conclusions and Future Directions**

The discussed experiments indicate that while motor factors, such as movement imitation, play a dominant role in shaping joint action performance (Kourtis et al., 2013; Atmaca et al., 2011; Sebanz et al., 2004, 2005), social factors like the partner's racial membership can also influence interactive performance, particularly in terms of reaction times. These findings align with previous research showing that implicit racial biases can affect social cognitive dynamics (e.g., Avenanti et al., 2010; Azevedo et al., 2013; Contreras-Huerta et al., 2013; Gutsell & Inzlicht, 2010; Sheng & Han, 2012), but they also extend previous works investigating these effects in the context of interactive motor tasks (Sacheli et al., 2015a).

Importantly, the neural underpinnings of those biases and their modulatory effects on motor behavior during joint actions need further investigation. The inconclusive evidence for an involvement of mPFC in racial bias expression and in motor coordination with racial out-group members, raises questions about the specific role played by this brain region in prejudice regulation. Moreover, while our findings suggest an influence of racial attitudes on interpersonal interactions, future studies should examine whether these effects could be mitigated by other factors, as for example familiarity with the interaction partner.

In conclusion, the present study shed light on the dynamic interplay between social and motor factors in joint action performance. It contributes to deeper understand the socio-cognitive processes involved in humans interactions and the potential barriers to efficient cooperation with an out-group.

## **2.5. Interim Summary**

The studies presented across Chapters 2 examined the influence of visuo-tactile integration, peripersonal space representation, body ownership and ethnic bias on interpersonal motor interactions. Chapter 2.1 investigates how the integration of visual and tactile information is modulated during joint actions. It demonstrates that PPS representations recruitment occurs not only when visual stimuli are aligned with one's own actions aim but also when they are associated with a partner's movements. This reflects a dynamic multisensory-motor interface during interpersonal interactions, where the partner's actions are integrated as extensions of one's own body and goal. The study, across three experiments, supports the idea that action prediction and motor coordination with others shape multisensory perception and expand the representation of peripersonal space toward the partner of an interaction. These findings highlight the critical role of sociality in modulating sensorimotor processes, particularly during joint actions where prediction and adaptation to a partner's movement is necessary to achieve a shared goal.

Chapter 2.2 builds on these findings by exploring cross-modal visuo-tactile interference effects during interpersonal interactions, both behaviorally and neurophysiologically. The study reveals that the integration of visuo-tactile and motor-proprioceptive information during joint tasks leads to improved performance in subsequent visuo-tactile tasks and early modulations in neural processing, depending on the type of interaction. This chapter extends individual cross-modal interference effects to an interpersonal context, showing that previous interactions shape both behavioral responses and neural patterns in the early stages of sensory processing. It suggests that interpersonal interactions induce lasting changes in sensory integration, depending on the multisensory demands of the task.

Chapter 2.3 shifts the focus to body ownership, using a novel approach to the Rubber Hand Illusion. The results indicate that stroking synchronicity reinforces the sense of embodiment and ownership of the other hand and touch when introducing an interpersonal interaction. The study highlights the dissociation between subjective embodiment, proprioception drift toward the other, and subsequent multisensory visuo-tactile interference effects, suggesting that social interactions can modulate body perception in ways that are distinct from classical RHI paradigms.

Chapter 2.4 explores how social categorization based on ethnicity affects motor coordination during joint actions, particularly examining the role of the mPFC on motor behaviors when interacting with in-group and out-group partners. Interestingly, participants showed no baseline racial bias, in line with previous research that often highlights that implicit biases are not always

significant (Cooley et al., 2018; Payne et al., 2005; Steele et al., 2018). The main results was on performance in the joint action task, RTs were found to be slower when interacting with out-group avatars, suggesting a subtle influence of social cognitive biases on sensorimotor processes. Overall, the results contribute to understanding the interplay between social factors and motor behaviours in joint actions, highlighting potential barriers to effective cooperation across different groups and expanding on previous findings by Sacheli et al. (2015b).

Together, these studies contribute to understand the role of social interactions dynamics in shaping multisensory integration, body representation and motor processes. They underscore the importance of prediction, coordination, and social contexts in extending low-level sensory-motor processes to higher-order cognitive dynamic in interactive settings, ultimately contributing to a more comprehensive understanding of human experience in social exchanges.

**3**

**Communication Dynamics in Interpersonal  
Contexts**

## **3.1. Breaking the ice in a conversation: abstract words prompt dialogs more easily than concrete ones.**

### **3.1.1. Introduction**

Abstract concepts (ACs) (i.e., “democracy,” “justice,” “peace”) are thought to be mainly acquired through linguistic interactions (Wauters et al., 2003) because they rely less on experiences made through the five senses than concrete concepts (CCs). Their meanings are grounded in social interactions, which represent the essential scaffolding allowing the emergence of “shared-cultural labels” conveying complex knowledge (Borghi, 2022; Fini et al., 2021, 2023; Fini & Borghi, 2019). Evidence suggesting the involvement of mouth motor areas during the processing of ACs supports this view (Borghi et al., 2011; Borghi & Zarccone, 2016; Dreyer & Pulvermüller, 2018; Fini et al., 2022; Ghio et al., 2013; Granito et al., 2015; Mazzuca et al., 2018). According to the Words As social Tool (WAT) proposal (Borghi et al., 2017, 2018, 2019), mouth activation may be interpreted as: (i) preparing to complement our knowledge by asking reliable sources [Social Metacognition (SM), Borghi, 2022; Borghi et al., 2017, 2018, 2019; Shea, 2018]; (ii) re-explaining to ourselves the meaning associated with the concept and (iii) simulating the linguistic experience through which we have learned ACs in the first place. The specific mechanisms – not mutually exclusive – that might activate mouth motor areas share a common ground: ACs evoke linguistic experiences.

Linguistic experiences represent the modality through which humans can efficiently communicate and create a space of intersubjective knowledge. Abstract knowledge emerged, indeed, as the result of intellectual agreements on shared meanings attributed to experiences and validated by communities (see Gilead et al., 2020). Although the process of conflating different experiences in a single symbol characterizes all concepts and words (e.g., Deacon, 1998; Dunbar, 1998), the heterogeneity of experiences that ACs evoke makes people more uncertain of their meaning and renders the input and interaction with others particularly crucial. During their use, we might need to rely on other people to understand the meaning of words on which we are uncertain or to negotiate their meaning (Borghi, 2022; Mazzuca & Santarelli, 2022). Consequently, the full development of abstract representations might rely upon successful linguistic exchanges among conspecifics (see Dunbar, 1998). ACs would not only be the “glue” holding together scattered and heterogeneous information (see the notion of situational systematicity in Davis et al., 2020), but they would also represent the “social glue” providing a common reference of knowledge within

societies (Borghi, 2022; Borghi & Tummolini, 2020; Mazzuca et al., 2021; Mazzuca & Santarelli, 2022). For example, values, cultural notions, scientific knowledge, and socio-political definitions are the result of a diachronic linguistic exchange which contributes to creating a connection among individuals. Along these lines, dialogical exchanges might be a source allowing us to detect the social nature of ACs. Because of the relevance of social interactions for ACs definition, recently authors have underlined the necessity to investigate ACs' use, for example, during conversations (Banks et al., 2022; Borghi, 2022), since they represent the spontaneous human practices of being in relation. Social interactions, thus, appear to be one of ACs' learning environment, and the question arises as to how spontaneous conversations use ACs (Borghi & Binkofski, 2014; Dove, 2009, 2011, 2014; Prinz, 2012). While many studies focused on semantic, emotional, social (Della Rosa et al., 2010; Diveica et al., 2022; Villani et al., 2019), and interoceptive conceptual dimensions (Connell & Lynott, 2012; Villani et al., 2021), more pragmatic dimensions of ACs still need to be explored (for an exception, see Villani et al., 2022). Pragmatic dimensions might offer important insights for accurately selecting stimuli materials in new interactive ecological paradigms. A promising research avenue, indeed, consists in studying ACs in natural environments where they are acquired and used for communication. The adoption of more ecological experimental settings in which dialogs spontaneously evolve might help to better understand i) the use of concepts in daily life and ii) the psychological and bodily synergies which are both causes and effects of linguistic exchanges. However, if, on the one hand, interactive paradigms represent a valuable and ecological experimental approach, on the other hand, they also require an additional effort to better control for unexplored variables which might intervene during linguistic exchanges. In this regard, here we explore a specific pragmatic dimension that we called "Easiness to Start a Conversation"(ESC), which touches upon aspects related to the capability of a concept to trigger a conversational exchange. Given their intrinsic social and dialogical nature, we expect that compared to CCs, ACs might more easily "afford" conversations.

In the current study, we operationalized this conjecture by investigating the perceived facility to start a conversation on concepts of different abstractness levels. To this aim, we asked participants to rate abstract and concrete words along different dimensions, spanning from more traditional to less explored ones. These are labeled here as semantic/metacognitive [i.e., SM – perceived need of others, Word Confidence (WC), Contextual Availability (CA)], communicative/pragmatic [i.e., Openness to Negotiation (ON), ESC), and emotional–experiential (i.e., Pleasantness, Valence, Familiarity). Specifically, ESC concerns the estimation of how easy it is to start a conversation with a specific concept. SM (Villani et al., 2019) refers to the degree to which the participant feels the

need to rely on others to understand the concept's meaning. WC (Mazzuca et al., 2022) relates to the extent to which participants feel confident in their understanding of a concept's meaning. ON refers to the degree to which participants feel prone to negotiate with others the concepts' meaning (see Mazzuca & Santarelli, 2022). CA (Schwanenflugel et al., 1988) refers to the ease with which the participant can think of a context or a circumstance associated with the word or in which the word may appear. Familiarity (FAM, Barca et al., 2002) measures how much participants think they know the word. Finally, Valence (Bradley & Lang, 1999; Mohammad, 2018) refers to the degree of positiveness/negativeness of a word, while Pleasantness relates to how much a concept is judged to be pleasant or unpleasant. In particular, we decided to focus here just on the ESC dimension because it represents a valuable aspect for studying the relationship between lexicon and social interaction. Gathering knowledge of the kinds of words considered more suitable to start a dialog is the very first step to approaching conversational dynamics. The other new pragmatic dimension ON, equally worth to be explored, will be investigated in more detail, in separate studies. Because the meaning of ACs finds its natural source in socio-linguistic exchanges, we predict that ACs will be rated higher on ESC than CCs, regardless of their scores on Pleasantness, Valence, and Familiarity. In line with previous evidence, here we also predict that ACs will be rated higher than CCs on ON because of the indeterminacy of their meaning (see Mazzuca & Santarelli, 2022). They should also score higher on SM because others are needed to understand or negotiate their meaning, and lower on WC. In line with the literature, ACs should also score lower in CA than CCs. In addition, we expect that higher ratings on ESC for ACs will be partially explained by higher ratings on SM, ON, WC, and CA.

Moreover, we explored whether different subclusters of abstract (Villani et al., 2019) and CCs differ in their rating scores for ESC and whether higher ratings on SM, ON, WC, and CA would partially explain these results. The philosophical–spiritual concepts (PS) (i.e., “prestige,” “paradise,” “infinity”) refer to religious words, principles, concepts linked to argumentation, reasoning, decision-making, and negatively connoted words, related to characteristics of the self. The emotional concepts (EM) (i.e., “pride,” “irony,” “patience”) refer to emotions, mental states, emotionally connoted social situations, and characteristics of the self with respect to others. Then, self-sociality concepts (SS) (i.e., “politeness,” “attention,” “merit”) indicate the interpersonal mental states and finally, the physical spatio-temporal and quantitative concepts (PSTQ) (i.e., “sum,” “number,” “subtraction”) refer to interactions with external objects/entities (Villani et al., 2019). We were interested in understanding whether these subclusters of concepts might be differently categorized as a function of the examined dimensions, in particular of the new

communicative/pragmatic dimensions. Finally, we performed an exploratory study involving an English native sample to gather knowledge about the cross-cultural evaluation of the same dimensions in AC and CC.

## Study 1

Study 1 hypotheses, methods, and analyses were formally pre-registered ([https:// osf.io/sym9p](https://osf.io/sym9p)). As declared in the preregistration, we performed a rating study with an Italian sample of participants to investigate the dimension of ESC along with the dimensions of Pleasantness, Valence, Familiarity, SM, ON, WC, and CA. Stimuli were both AC and CC.

## Method

### *Participants*

A total of 48 Italian participants took part in the study in a window between January 31 and April 8 2022. Participants were recruited via anonymous links either by posting the surveys on social networks (Facebook, Twitter) or spreading the questionnaires through the research team's extended network of acquaintances. Italian was the native language of all the participants, who were naïve about the purpose of the experiment. Their participation was voluntary and not reimbursed. The required sample size was previously estimated through a statistical power analysis (Cohen, 2013), performed with the software More Power 6.0.4 (Campbell & Thompson, 2012) based on the effect size obtained in a preliminary pilot study, as described in the preregistration (<https://osf.io/sym9p>). More specifically, in the pilot study, we asked 115 participants in 4 different surveys to rate 108 words (68 abstract, 60 concrete) on Pleasantness, Valence, and ESC. It emerged that ACs were rated significantly higher ( $M = 4.06$ ;  $SE = 0.12$ ,  $\eta^2 = 0.38$ ) on ESC dimension, as compared to CCs ( $M = 3.11$ ;  $SE = 0.13$ ). Thus, in Study 1 we used the partial eta squared value (0.38), that is, the effect size of the main effect of Category of Concepts (abstract, concrete) on the dependent variable ESC in the pilot study as the expected effect size for the power analysis. The output indicated that for a two within-factors design (Category, abstract vs. concrete), 24 participants rating each dimension were sufficient to reach a power of 0.95 and an eta squared of 0.38.

### **Materials and procedure**

To prevent participants from fatiguing effects, we split the sample of words into two different surveys. The first survey was composed of two randomly presented blocks where 108 words were

rated on Valence, ESC, Pleasantness, and Familiarity (Block 1), and CA, SM, WC, and ON (Block 2). The second survey was structured as the first survey, but it featured an additional sample of 108 words. Since the surveys to be compiled were two, by taking into account the indication gathered from the power analysis, a sample of 48 participants was needed. A total of 144 participants were originally contacted. We excluded 96 participants, mostly because they did not complete the surveys, likely due to the surveys length ( $N = 92$ , 64.33% of the original sample), and a few because they provided random responses and faster than 10 minutes ( $N = 4$ , 0.02% of the sample). This resulted in a final sample of 48 participants, among which 24 participants (17 females and 7 males;  $M$  age = 24.88 years,  $SD = 3.05$ ) rated the first subset of 108 words and the remaining 24 participants (17 females and 7 males; mean  $M$  age = 26.58 years,  $SD = 9.05$ ) rated the second subset of 108 different words. Ethics permission was granted by the Ethics Committee of the Psychology Department at Sapienza University of Rome (Prot. n. 0001040–16/11/2020). Participants were informed of the general purpose of the study. Each survey provided informed consent. The study was implemented in Psytoolkit (Stoet, 2010, 2017), a Interactive-to-use toolkit for experiments and survey design. The stimuli were 108 abstract and 108 concrete words. The number of items was defined on the basis of the analyses conducted on a preliminary data collection described in the preregistered report (<https://osf.io/sym9p>). We doubled the number of items of the preliminary data collection in order to increase analysis reliability. ACs selected from Villani et al.'s (2019) database were counterbalanced for the four subclusters identified in Villani et al.'s (2019) and randomly assigned to one of the two surveys. Accordingly, we selected 27 physical, spatio-temporal, and quantitative concepts, for example, “scheme”; 27 philosophical and spiritual concepts, for example, “judgment”; 27 concepts related to the self and social concepts, for example, “politeness”; and 27 emotional and inner state concepts, for example, “happiness.” CCs were selected from Montefinese et al. (2014) database and sampled to include the types of CCs more frequently identified in the literature (see Rumiati & Foroni, 2016; Villani et al., 2021; Warrington & Shallice, 1984). Specifically, we selected 36 food concepts, 36 animal concepts, and 36 artifacts. Participants were asked to rate each concept on eight dimensions, that is, Valence, Pleasantness, ESC, SM, WC, Context Availability, Familiarity, and ON through a Likert scale ranging from 0: “completely negative” to 7: “completely positive.” We asked participants to evaluate some Valence of each word, the SM dimension, that is, the need to rely on other to better understand a concept; the Familiarity of each word; the WC, that is the degree to which we think to know the meaning of a word and, and the CA, that is, the degree to which we perceive easy to think about a context for each word. Moreover, we included three new dimensions: the ON dimension, referring to how

much we feel open to negotiating the meaning of words; the Pleasantness of a word; and finally, the ESC, which is the specific object of investigation in the current study. ESC refers to how much we consider easy to use a word to start a dialog (for further information about the scales, see next section). Once participants agreed to take part in the study, they were presented with the rating task and asked to evaluate target words on the eight dimensions using the Likert scales scale. Before rating the concepts, participants were asked to report their age in numbers and their gender. They gave their responses into separate boxes for each target word presented per dimension. All data, scripts, and analyses are available at <https://osf.io/tq3nb>. Each participant was randomly assigned to compile one of the two surveys, hence rating a single subset of 108 words on all the dimensions across two blocks. The order of blocks and dimensions was randomized across participants, and items were randomly presented within each block.

#### *Description of the scales administered in the study*

Specifically, for the Valence dimension, the instructions were: ‘Please, rate the valence of each of the following words from 0: ‘completely negative’, to 7: ‘completely positive’; for the Social Metacognition dimension, the instructions were: ‘Please, evaluate the extent to which you perceive or have perceived a need to rely on others in order to understand the meaning of each of the following words’, from 0: ‘no need to rely on other’ to 7: ‘completely need to rely on other’. For the Familiarity dimension, the instructions were: ‘Please, rate the extent to which you perceive as familiar each of the following words’, from 0: ‘not familiar at all’ to 7: ‘completely familiar’. For the Word Confidence dimension, the instructions were: ‘Please, evaluate the degree to which you think to know the meaning or to understand the functioning of the concept to which each of the following words refer’, from 0: ‘not at all’ to 7: ‘completely’. For Context Availability, the instructions were: ‘Please, rate the degree to which you perceive easy to think about a context for each of the following words’ from 0: ‘not easy at all’ to 7: ‘completely easy’. Importantly, Pleasantness, Easiness to Start a Conversation, and Openness to Negotiation were the new explored dimensions.

For the Pleasantness dimension, the instructions were: ‘Please, evaluate how pleasant you perceive each of the following words from 0: ‘not pleasant at all’, to 7: ‘completely positive’’. For the Easiness to Start a Conversation dimension, the instructions were: ‘Please, rate the easiness to start a conversation from each of the following word, 0: ‘not easy at all’ to 7: ‘completely easy’. Finally, for the Openness to Negotiation dimension, the instructions were: ‘Please, evaluate how much you

feel open to negotiate with others the meaning of each of the following words', from 0 'not open at all to negotiate the meaning' to 7: 'completely open to negotiating the meaning'.

## Data analysis

Data pre-processing and data analysis were performed with R (R-Core Team, 2019) and RStudio (version 4.2.0). Statistical significance of fixed effects for mixed models was determined through type III ANOVA test from "afex" package (Singmann et al., 2015), and p-values were calculated with Satterthwaite's approximation. Post hoc comparisons were performed with the "emmeans" package (Lenth et al., 2019) using Tukey's correction for multiple comparisons. Principal component analysis (PCA) was carried out with "tidymodels"(Kuhn & Wickham, 2018) and hierarchical cluster analysis with "FactoMineR" (Lê et al., 2008).

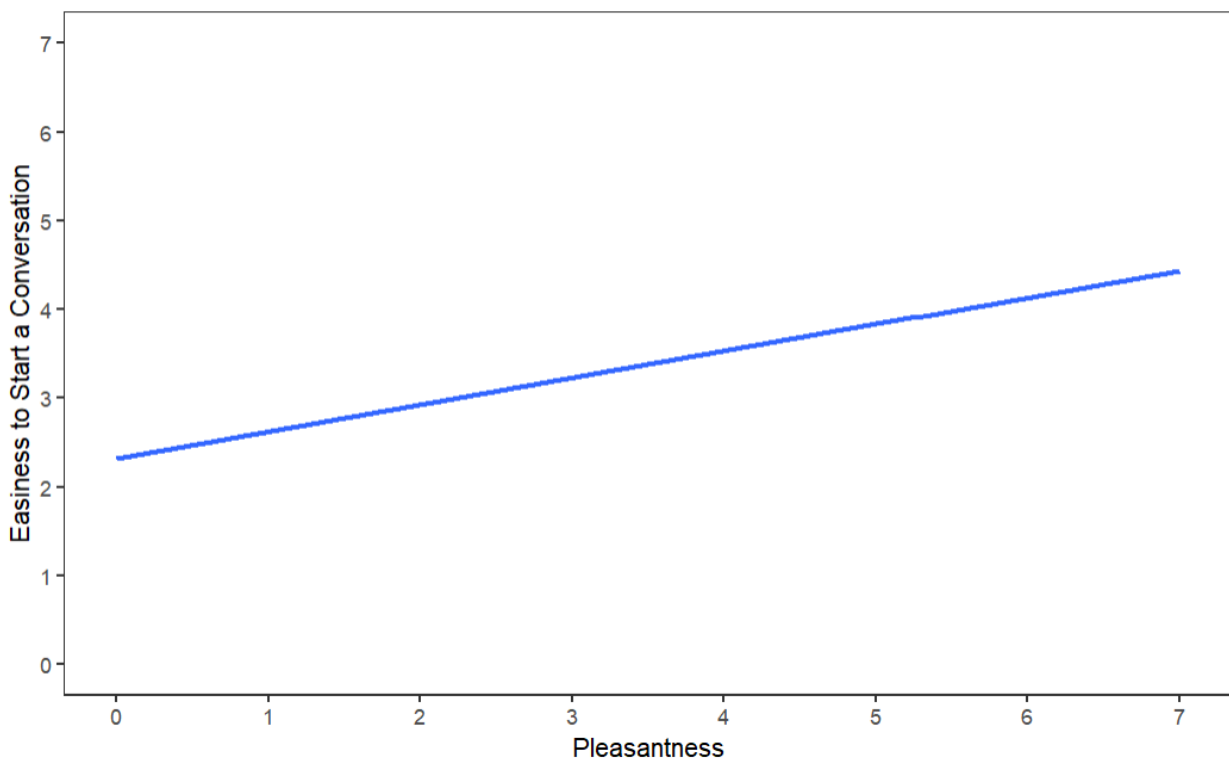
## Results

### *ACs are rated higher on ESC compared to CCs*

In order to assess whether ACs would be rated higher on ESC compared to CCs regardless of their scores on Pleasantness (ACs  $M = 3.7889$ ,  $SD = 1.9778$ ; CCs  $M = 3.7527$ ,  $SD = 1.7609$ ), Valence (ACs  $M = 4.0945$ ,  $SD = 2.1590$ ; CCs  $M = 4.0308$ ,  $SD = 1.9073$ ), and Familiarity (ACs  $M = 5.3912$ ,  $SD = 1.8188$ ; CCs  $M = 5.6878$ ,  $SD = 1.7315$ ), we used linear mixed-effects models (Bates et al., 2014). In the first model ( $R^2C = 0.31$ ), we entered ESC scores as dependent variable and Category of Concepts (abstract vs. concrete) as a fixed effect. We included participants and words as random intercepts. The model showed a main effect of Category of Concepts,  $F(1,214) = 14.65$ ,  $p = 0.0001$ , with higher means in the Abstract compared to the CCs condition (ACs  $M = 3.65$ ,  $SE = 0.191$ ; CCs  $M = 3.27$ ,  $SE = 0.191$ ). Our second aim was to investigate whether this effect was modulated by Pleasantness, Valence, and Familiarity (i.e., the emotional-experiential dimension). To this end, we fitted additional models, which we describe below. To avoid multicollinearity, we computed Pearson correlations among these variables, which showed that Valence and Pleasantness were highly correlated,  $R = 0.66$ ,  $p < 0.0001$ . Familiarity was correlated with Valence,  $R = 0.42$ ,  $p < 0.03$ , but not with Pleasantness,  $R = 0.24$ ,  $p = 0.26$ . Based on these results, we fitted two different models, both modeling ESC scores as a function of Category of Concepts (abstract vs. concrete) and including participants and items as random intercepts but varying for the chosen continuous predictors. Specifically, the first model (M1) included Familiarity and Pleasantness as continuous predictors, while the second model (M2) included Valence as a continuous predictor.

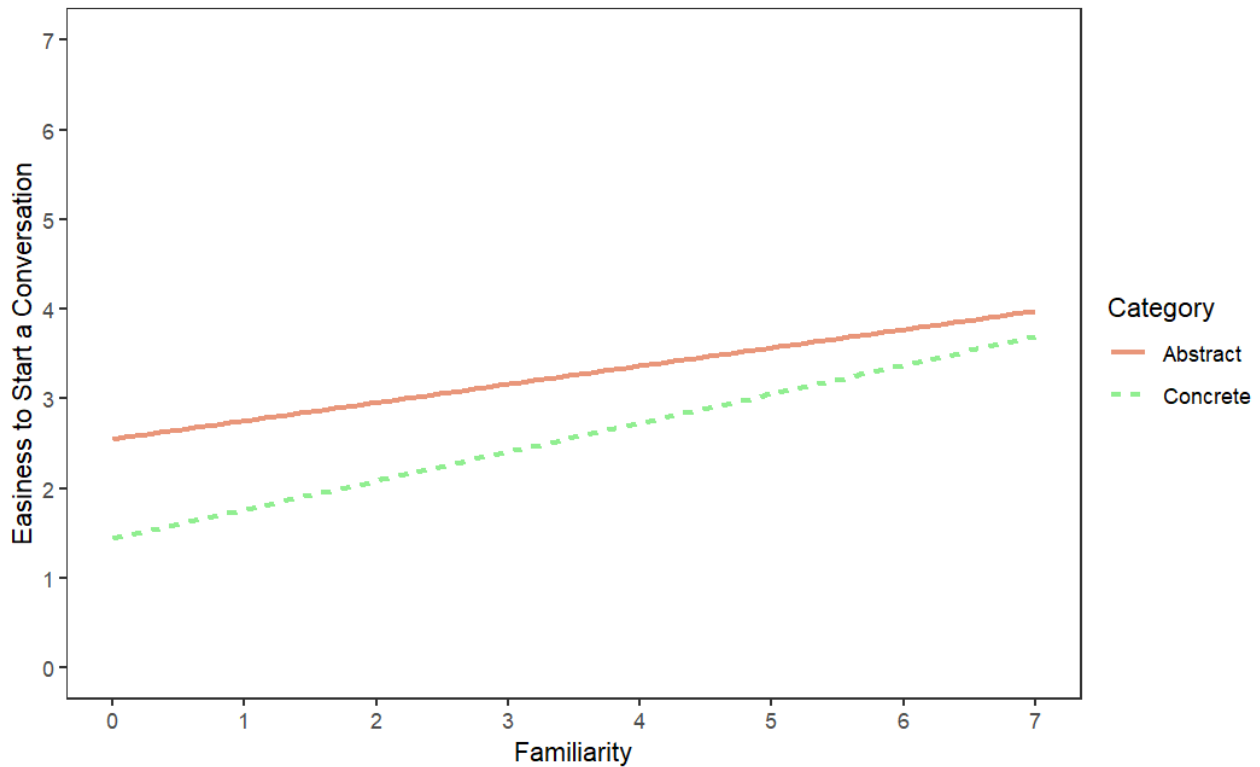
## MI

The model with Pleasantness and Familiarity as continuous predictors ( $R^2C = 0.32$ ) yielded a significant main effect of Category of Concepts,  $F(1,210.1) = 26.972$ ,  $p < 0.0001$ , Familiarity,  $F(1,5124.1) = 128.851$ ,  $p < 0.0001$ , and Pleasantness,  $F(1,4365.6) = 169.137$ ,  $p < 0.0001$ . The higher the ratings of Pleasantness, the higher the ratings of ESC (see Figure S1). The main effects of Category of Concepts and Familiarity were better explained by a significant two-way interaction of Familiarity x Category of Concepts,  $F(1,5119.7) = 25.696$ ,  $p < 0.0001$ . Simple slope analysis revealed that the slopes of AC (LCI 0.0794 – UCI 0.158) and CC (LCI 0.2086 – UCI 0.293) were significantly different from zero as a function of the factor Familiarity. The pairwise difference between the simple slopes of AC and CC as a function of Familiarity was significant, estimate = 0.132, SE = 0.026,  $t(5121) = 5.068$ ,  $p < 0.0001$ , showing that more familiar concepts are also the ones with which it is easier to get a conversation started.



**Figure S1.** The graph shows the predicted values of the outcome variables. Shaded bands represent the confidence intervals (95%). The higher participants' scores of Pleasantness, the higher the Easiness to Start a Conversation scores.

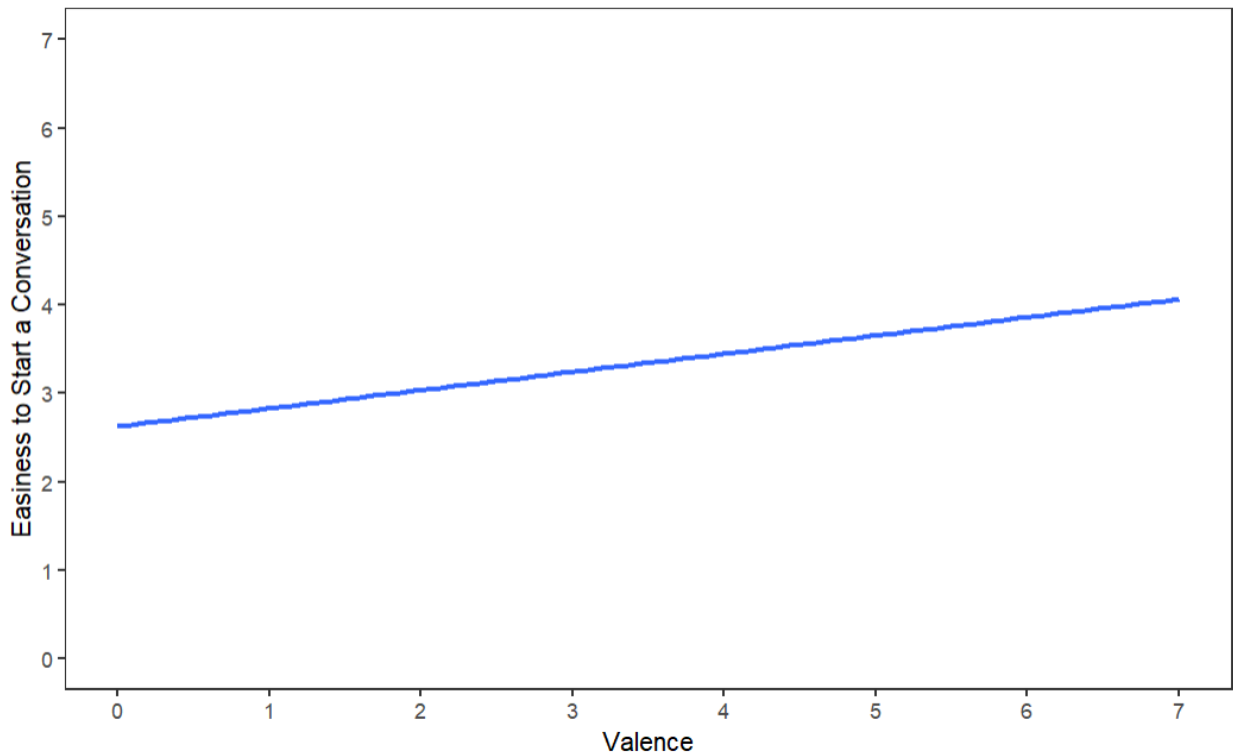
It is worth noting that this also mediated the advantage of ACs over CCs previously discussed (see Figure S2). More specifically, the lower the Familiarity, the higher the difference between ACs and CCs.



**Figure S2.** The graph shows the predicted values of the outcome variables. Shaded bands represent the confidence intervals (95%). The higher participants' scores of Familiarity, the higher the Easiness to Start a Conversation scores.

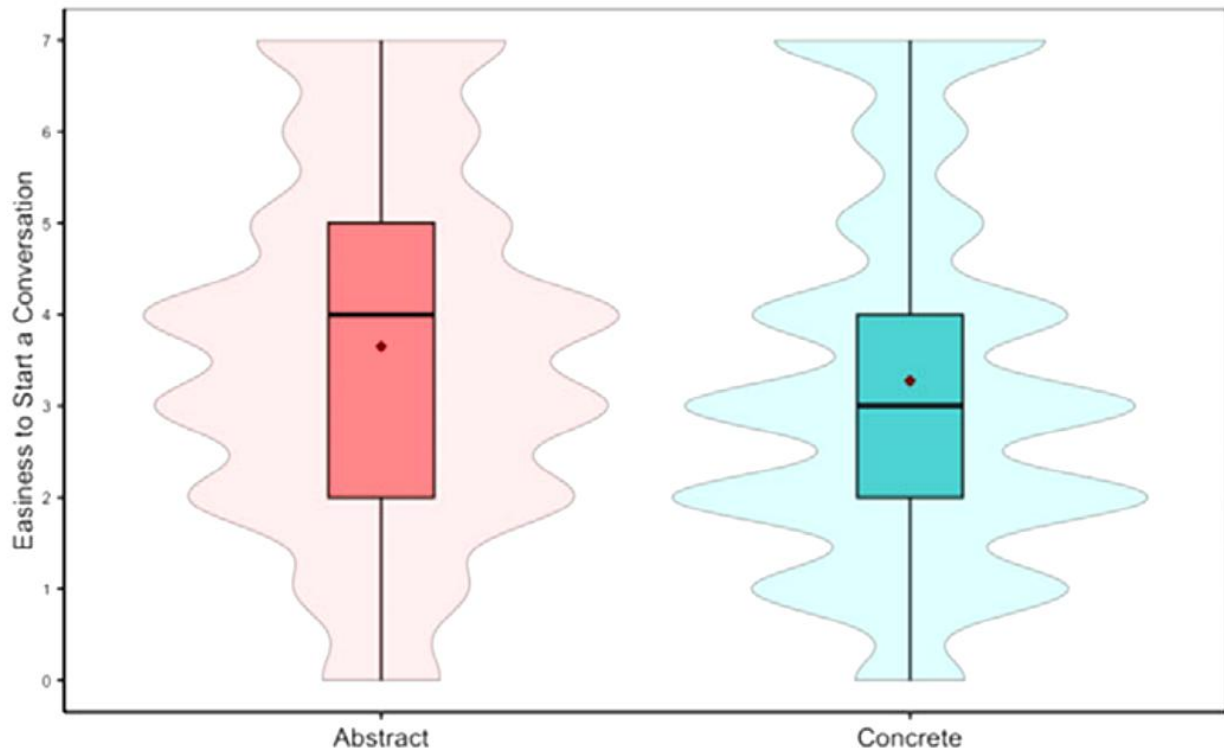
## M2

The model with Valence as a continuous predictor ( $R^2 C = 0.31$ ) yielded a main effect of Category of Concepts,  $F(1,213.4) = 15.5628$ ,  $p < 0.0001$ , and Valence,  $F(1,3760.6) = 133.9786$ , while the two-way interaction of Category of Concepts x Valence was not significant  $F(1, 4160.4) = 0.2745$ ,  $p = 0.60$ . More valenced concepts are also those with which it is easier to start a conversation (see Figure S3).



**Figure S3:** The graph shows the predicted values of the outcome variables. Shaded bands represent the confidence intervals (95%). The higher participants' scores of Valence, the higher the Easiness to Start a Conversation scores.

In conclusion, from the analysis emerged that ACs are rated higher on ESC compared to CCs (Fig. 1). Moreover, at the increase of concepts' Familiarity, the ACs advantage over CCs on ESC disappears. Finally, regardless of conceptual category (abstract/concrete), the more pleasant and positive valenced concepts are evaluated, the higher their ESC score.

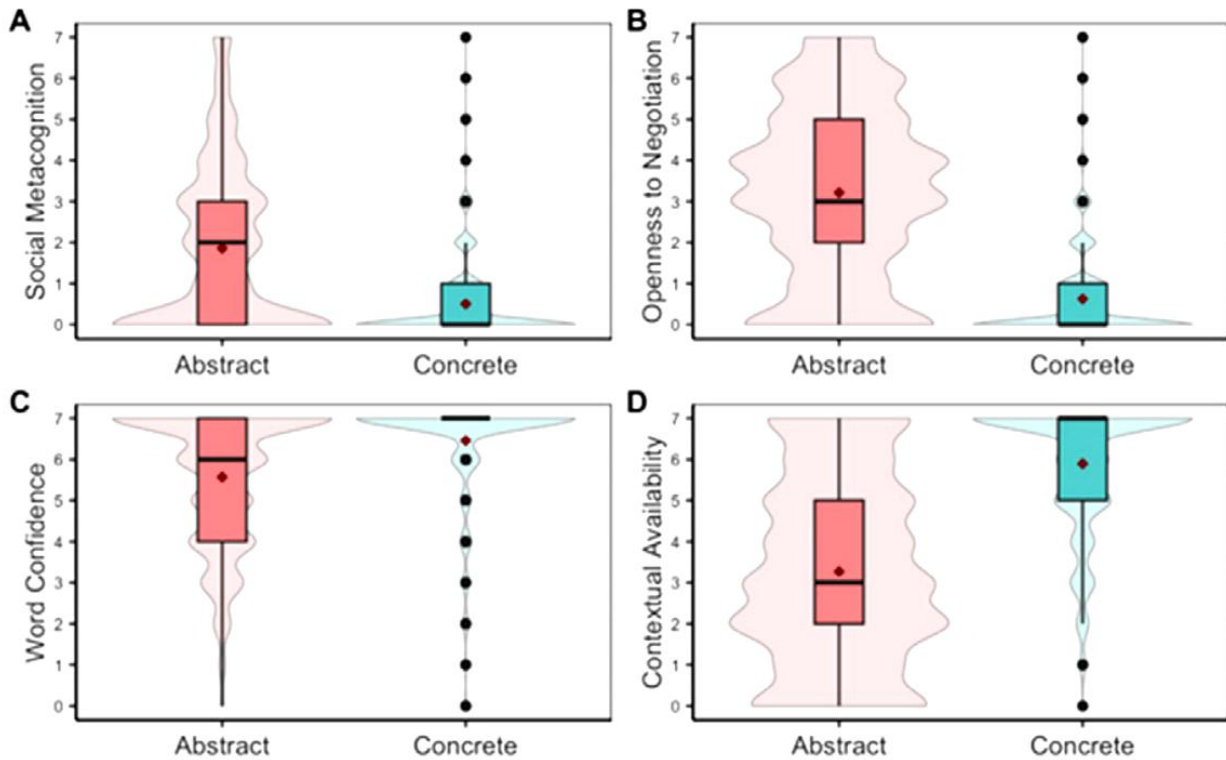


**Fig. 1.** Abstract concepts (ACs) are rated significantly higher than concrete concepts (CCs) in the Easiness to Start a Conversation (ESC). Horizontal lines in the boxes indicate the median, upper and lower borders indicate first and third quartiles, red rhombus represents the average of observations, and “whiskers” extend to the farthest points that are not outliers.

*ACs are rated higher on SM, ON, and lower on WC and CA compared to CCs*

In keeping with our first prediction, we explored whether AC and CC differed in their ratings on SM, ON, WC, and CA using four linear mixed-effects models. Each of the models included Category of Concepts (abstract vs. concrete) as fixed factor, participants and items as random intercepts, and varied for the dependent variable used. The model with SM as dependent variable ( $R^2C = 0.42$ ) showed a main effect of Category of Concepts  $F(1,214) = 555.94, p < 0.0001$  (ACs  $M = 1.855, SE = 0.176$ ; CCs  $M = 0.504, SE = 0.176$ ), with higher means in ACs compared to CCs condition (see Fig. 2A). The model with ON as dependent variable ( $R^2C = 0.52$ ) showed a main effect of Category of Concepts  $F(1,214) = 852.25, p < 0.0001$  (ACs  $M = 3.21, SE = 0.142$ ; CCs  $M = 0.626, SE = 0.142$ ), with higher means in ACs compared to CCs condition (see Fig. 2B). The model with WC as dependent variable ( $R^2C = 0.34$ ) showed a main effect of Category of Concepts  $F(1,214) = 271.72, p < 0.0001$  (ACs  $M = 5.57, SE = 0.151$ ; CCs  $M = 6.45, SE = 0.151$ ), with lower means in ACs compared to CCs condition (see Fig. 2C). The model with CA as dependent variable ( $R^2C = 0.49$ ) showed a main effect of Category of Concepts  $F(1,214) = 784.43, p < 0.0001$  (ACs  $M$

= 3.27, SE = 0.171; CCs M = 5.89, SE = 0.171), with lower means in ACs compared to CCs condition (see Fig. 2D).



**Fig. 2.** The plots show abstract concepts (ACs) and concrete concepts (CCs) scores on Social Metacognition (SM), Openness to Negotiation (ON), Word Confidence (WC), and Contextual Availability (CA). Horizontal lines in the boxes indicate the median, upper and lower borders indicate first and third quartiles, red rhombus represents the average of observations, and “whiskers” extend to the farthest points that are not outliers. **Panel (A)** ACs (M = 1.855, SE = 0.176) are rated higher on SM than CCs (M = 0.504, SE = 0.176). **Panel (B)** ACs (M = 3.21, SE = 0.142) are rated higher on ON than CCs (M = 0.626, SE = 0.142). **Panel (C)** ACs (M = 5.57, SE = 0.151) are rated lower on WC than CCs (M = 6.45, SE = 0.151). **Panel (D)** ACs (M = 3.27, SE = 0.171) are rated lower on CA as compared to CCs (M = 5.89, SE = 0.171).

In conclusion, results indicate that ACs are rated higher on SM, ON, and lower on WC and CA compared to CCs.

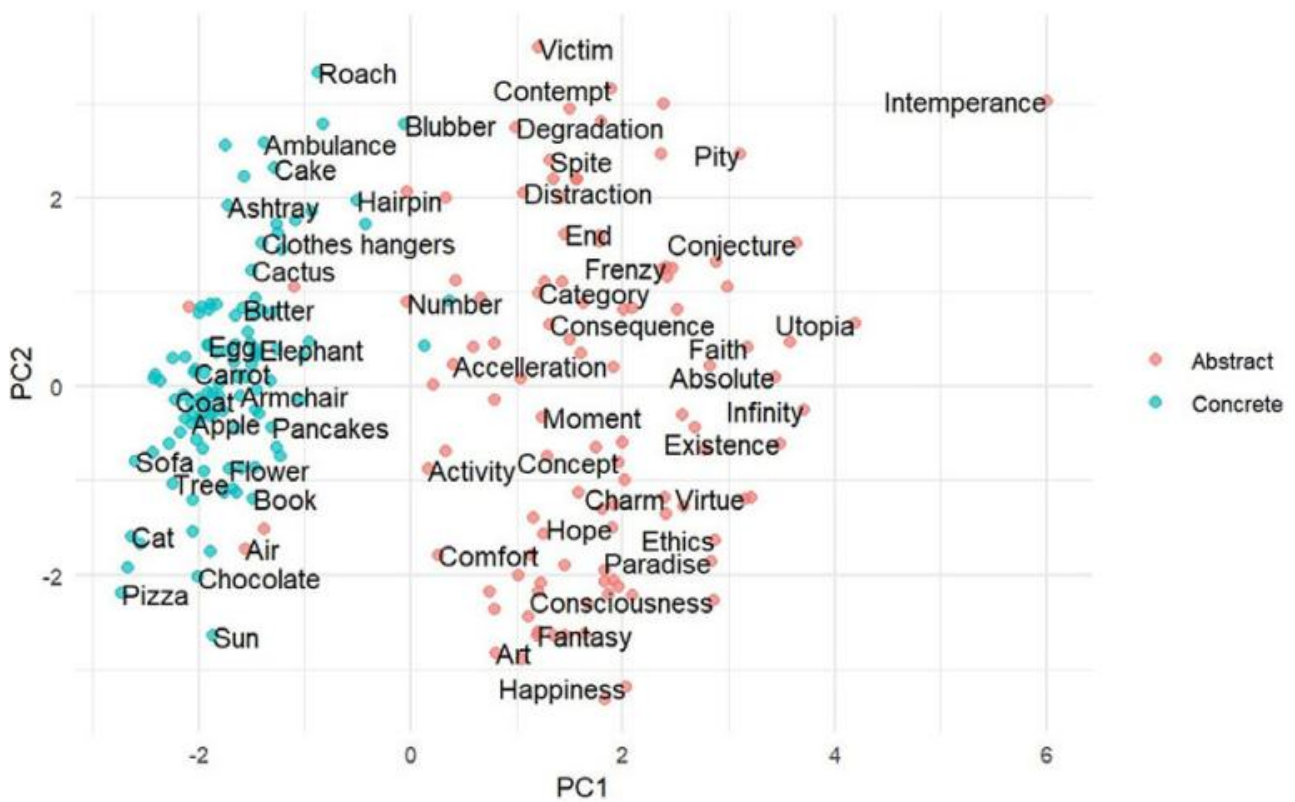
#### *Metacognitive–semantic and emotional–experiential components explain most of words variance*

Once we identified the dimensions that mostly explain ESC, we sought to explore how the words of our sample were distributed in the semantic space composed of our variables of interest as a function of ESC. So, we performed a PCA (Jolliffe, 2010) on the normalized rating dataset. PCA is a suitable method to extract a new set of features from an existing set of variables, explaining the variance in the dataset. Before applying PCA, all the variables (i.e., the average rating score per

each dimension) were centered and standardized. We entered ESC as outcome variable and the other dimensions (SM, ON, WC, and CA) as numerical predictors. We focused on components that weigh more of  $|\lambda|$  (Mazzuca et al., 2022; Villani et al., 2019). The first two components extracted by the PCA explained together 83% of variance, with Component 1 explaining 52% of variance and Component 2 explaining the remaining 31%. The first component was mostly composed of semantic and experiential dimensions: SM, ON, WC, and CA. The second component was instead mostly characterized by emotional–experiential dimensions: Pleasantness, Valence, and Familiarity (see Table 1). The first component was characterized by the opposition between SM and ON, WC, CA, and Familiarity (Fig. 3). This suggests that the more we are confident about the meaning of the word (WC), the more we can retrieve contexts evoked by words (CA); the less we need others to master the meaning of the word (SM), and the less we are open to negotiating the word’s meaning (ON). Crucially, AC and CC were sharply distinguished on the basis of the dimensions included in PC1, with ACs characterized by high scores on SM and ON and CCs characterized by high scores on WC and CA. The composition of the second component instead points to the fact that the more positively valenced, pleasant, and familiar the words are, the easier it is to start a conversation prompted by these words. In summary, through the (PCA) multivariate analysis, we identified two main components: a semantic–metacognitive one, composed of WC, SM, ON, and CA components and an emotional–experiential one composed of Pleasantness, Familiarity and Valence components. While the first one fully captures the abstract–concrete categorical distinction, the second one indicates that regardless of conceptual category, the concepts evaluated as more positive, pleasant, and familiar are also considered as being easier to start a conversation with.

**Table 1.** Contribution of each dimension on PC1 and PC2, and their positive and negative values

Dimensions	Values	Components
Familiarity	-0.28833	PC1
Pleasantness	-0.02321	PC1
Valence	-0.03107	PC1
Social metacognition	0.475916	PC1
Openness to negotiation	0.474826	PC1
Contextual availability	-0.47635	PC1
Word confidence	-0.48631	PC1
Familiarity	-0.37252	PC2
Pleasantness	-0.64173	PC2
Valence	-0.63348	PC2
Social metacognition	-0.11295	PC2
Openness to negotiation	-0.18734	PC2
Contextual availability	0.01059	PC2
Word confidence	-0.11086	PC2



**Fig. 3.** Although abstractness~concreteness was not considered as a dimension in the PCA, looking at the plot, we notice that abstract and concrete words are quite sharply distinguished on the basis of PC1. For instance, words like “knowledge,” “ethic,” and “virtue” are characterized by high scores on Social Metacognition and Openness to Negotiation, and low scores in Word Confidence and Contextual Availability, and words like “orange,” “sofa,” and “pencil” by the opposite pattern.

*The emotional–experiential component better predicts ESC as a function of abstractness*

After investigating how words are distributed in the semantic space composed of our targeted dimensions, we sought to determine which components contributed more to predict ESC. So, we fitted a linear model with ESC ratings as dependent variable, and the outcome of the PCA for each component as predictor ( $R^2C = 0.31$ ). The results showed a main effect of the second (PC2: emotional–experiential) component  $F(1,213) = 93.78, p < 0.0001$ . This shows that overall, emotional and experiential variables taken together predict ESC. Specifically, the more words were perceived as pleasant, positive, and familiar, the easier it is to start a conversation with them. Given our hypotheses, we were also interested in assessing whether there is a difference between AC and CC. So, we fitted a linear mixed model with ESC as dependent variable, Category of Concepts (abstract vs. concrete) as categorical predictor, the components as continuous predictors, and random intercepts for words and participants. The model ( $R^2C = 0.30$ ) yielded a main effect of the Category of Concepts  $F(1,215.1) = 5.41, p = 0.02$ , and of the second (PC2: emotional–experiential) component  $F(1,215.1) = 72.69, p < 0.0001$ . These two main effects were better explained by the significant two-way Category of Concepts x PC2 (emotional–experiential component) interaction  $F(1,215.1) = 8.60, p = 0.003$ . Simple slope analysis revealed that the slopes of AC (LCI 0.130 – UCI 0.258) and CC (LCI 0.275 – UCI 0.521) were significantly different from zero as a function of PC2. The pairwise difference between the simple slopes of AC and CC as a function of PC2 was significant (estimate = 0.204, SE = 0.0703,  $t(220) = 2.899, p = 0.0041$ ), showing that the more concepts are familiar, positively valenced, and pleasant, the higher the ESC and the less is the difference between AC over CC. In conclusion, the results indicate that PC2 (emotional–experiential component) is stronger as compared with PC1 (semantic–metacognitive component) in predicting the ESC dimension.

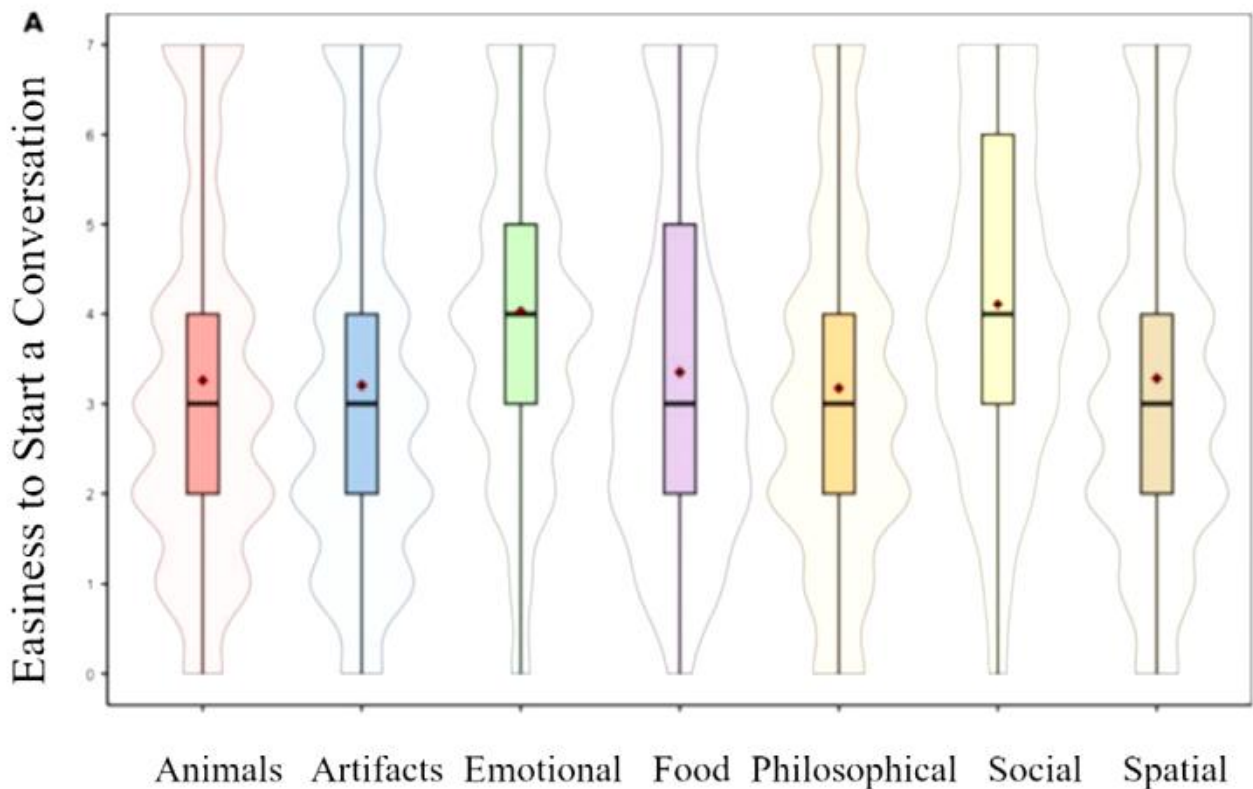
## **Explorative analysis**

*Self-Sociality and emotional concepts are rated higher on ESC compared to both other sub-kinds of ACs, and sub-kinds of CCs*

Recently, multiple representation theories have consolidated the idea that the abstract–concrete dichotomy does not suffice for the full identification of all the conceptual features. Many dimensions like interoception, emotions, language, and social interaction contribute to provide a more exhaustive conceptual clustering (see Banks et al., 2022). For this reason, we decided to explore whether specific ACs’ semantic subclusters identified by Villani et al. (2019) differ on the ESC dimension. To this end, we fitted a model (M1) with ESC as dependent variable, Subcluster of

Concept as a fixed factor, and random intercepts for participants and words ( $R^2C = 0.31$ ). 6.2. M1

The model showed a main effect of Subcluster of Concept  $F(6,209) = 10.366$ ,  $p < 0.0001$ . Tukey post hoc comparisons indicated that Emotional concepts ( $M = 4.03$ ,  $SE = 0.219$ ) were rated significantly higher on ESC compared to Food concepts ( $M = 3.35$ ,  $SE = 0.209$ ,  $t(209) = 4.011$ ,  $p < 0.0016$ ), Artifacts concepts ( $M = 3.21$ ,  $SE = 0.209$ ,  $t(209) = 4.869$ ,  $p < 0.0001$ ), Animals concepts ( $M = 3.26$ ,  $SE = 0.209$ ,  $t(209) = 4.546$ ,  $p = 0.0002$ ), Physical Space Time and Quantity concepts ( $M = 3.28$ ,  $SE = 0.219$ ,  $t(209) = 4.135$ ,  $p < 0.0001$ ), and Philosophical and Spiritual concepts ( $M = 3.18$ ,  $SE = 0.219$ ,  $t(209) = 4.726$ ,  $p < 0.0001$ ). Likewise, Self-Sociality ( $M = 4.11$ ,  $SE = 0.219$ ) concepts were rated significantly higher on ESC compared to all CCs, that is, Food concepts ( $M = 3.35$ ,  $SE = 0.209$ ,  $t(209) = 4.496$ ,  $p = 0.0002$ ), Artifacts concepts ( $M = 3.21$ ,  $SE = 0.209$ ,  $t(209) = 5.354$ ,  $p < 0.0001$ ), and Animals concepts ( $M = 3.26$ ,  $SE = 0.209$ ,  $t(209) = 5.032$ ,  $p < 0.0001$ ), and to the other abstract Physical Space Time and Quantity concepts ( $M = 3.28$ ,  $SE = 0.219$ ,  $t(209) = 4.589$ ,  $p < 0.0002$ ) and Philosophical and Spiritual concepts ( $M = 3.18$ ,  $SE = 0.219$ ,  $t(209) = 5.180$ ,  $p < 0.0001$ ). There was instead no difference on scores of ESC between Emotional and Self-Sociality concepts,  $t(209) = 0.454$ ,  $p = 0.999$  (see Fig. 4A). In summary, Emotional and Self-Sociality concepts scored significantly higher on ESC. Since the pattern of results of previous analyses suggested that Familiarity and Pleasantness significantly modulated ESC with AC and CC, we aimed at assessing whether this was the case also with more fine-grained conceptual categories. Specifically, we wanted to verify whether the advantage of Emotional and Self-Sociality concepts was due to higher scores in Familiarity and Pleasantness. So, we performed a second model (M2) ( $R^2C = 0.33$ ) with ESC as dependent variable, the Subcluster of Concepts as a fixed factor, as continuous predictors Familiarity and Pleasantness, and as random intercepts participants and words.

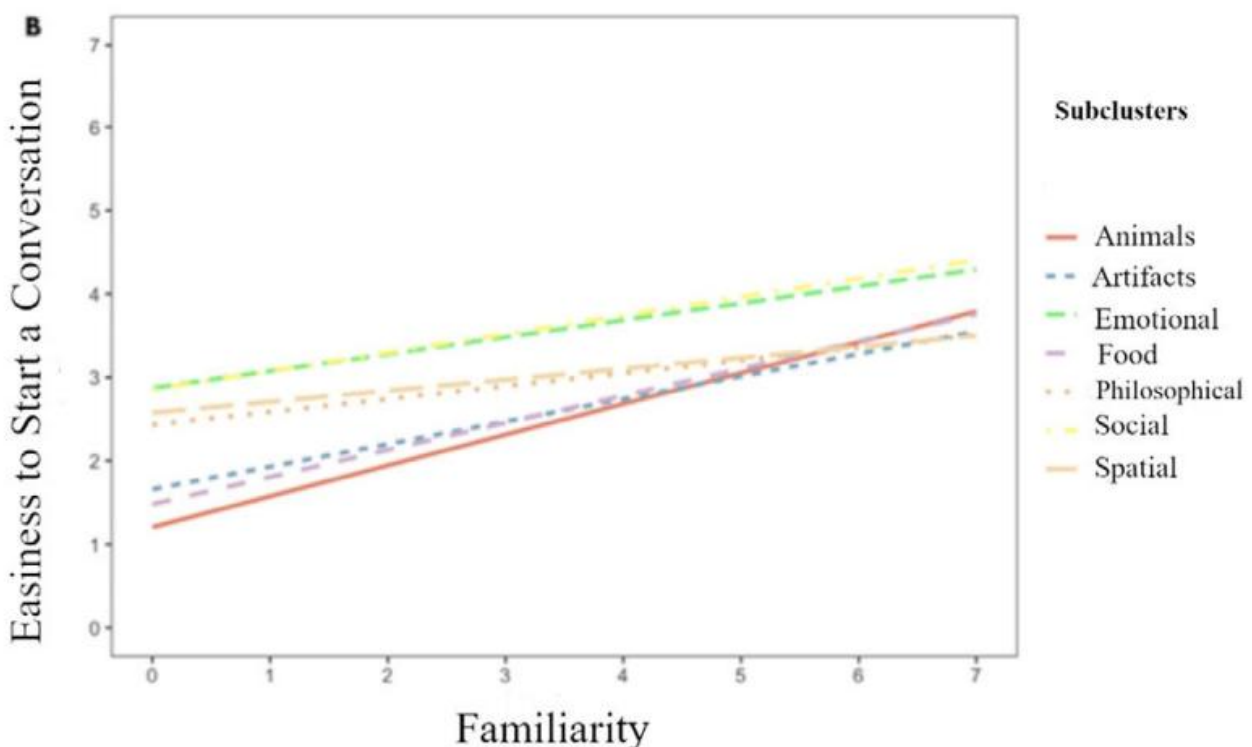


**Fig. 4A.** The plot shows the main effect of Subcluster of Concepts yielded by the model with Easiness to Start a Conversation (ESC) as dependent variable, Subcluster of Concepts as fixed factor, and participants and words as random intercepts, Emotional ( $M = 4.03$ ,  $SE = 0.219$ ) and Self-Sociality ( $M = 4.11$ ,  $SE = 0.219$ ) concepts scored significantly higher on ESC as compared to all the other Subclusters (Food:  $M = 3.35$ ,  $SE = 0.209$ ; Artifacts:  $M = 3.21$ ,  $SE = 0.209$ ; Animals:  $M = 3.26$ ,  $SE = 0.209$ ; Physical Space Time and Quantity:  $M = 3.28$ ,  $SE = 0.219$ ; Philosophical and Spiritual:  $M = 3.18$ ,  $SE = 0.219$ ).

## M2

The model showed a main effect of Subcluster  $F(6, 2479.9) = 8.6807$ ,  $p < 0.0001$ , of Familiarity  $F(1,5090.6)=111.7541$ ,  $p < 0.0001$ . The main effects of Familiarity and Subcluster were better explained by the two-way interaction Familiarity x Subcluster  $F(6, 5081.9) = 6.0737$ ,  $p < 0.0001$ . Simple slope analysis revealed that the slopes of all the subclusters were significantly different from zero as a function of Familiarity: Emotional (LCI 0.0866 – UCI 0.254), Artifacts (LCI 0.1560 – UCI 0.289), Animals (LCI 0.2278 – UCI 0.357), Food (LCI 0.1729 – UCI 0.310), Philosophical and Spiritual (LCI 0.0196 – UCI 0.146), Self-Sociality (LCI 0.053 – UCI 0.215), and Physical Space Time and Quantity concepts (LCI 0.0150 – UCI 0.154). The pairwise difference between the simple slopes of Animals Subcluster and Philosophical and Spiritual (estimate = 0.20960,  $SE = 0.0444$ ,  $t(5113) = 4.723$ ,  $p < 0.0001$ ), Physical Space Time and Quantity (estimate = 0.20812,  $SE = 0.0464$ ,  $t(5104) = 4.486$ ,  $p < 0.0002$ ), and Self-Sociality (estimate = 0.15823,  $SE = 0.0507$ ,  $t(5113)$

= 3.123,  $p = 0.0298$ ) Subclusters as a function of Familiarity was significant. The pairwise difference between the simple slopes of Artifacts Subcluster and Philosophical and Spiritual Subcluster was significant (estimate = 0.13944, SE = 0.0451,  $t(5087) = 3.092$ ,  $p = 0.0320$ ). The pairwise difference between the simple slopes of Food Subcluster and Philosophical and Spiritual (estimate=0.15874, SE=0.0458,  $t(5078)=3.469$ ,  $p < 0.001$ ) and Physical Space Time and Quantity (estimate = 0.15726, SE = 0.0477,  $t(5073) = 3.300$ ,  $p = 0.0169$ ) Subclusters was significant.



**Fig. 4B.** The plot shows the predicted values of the outcome variables. Shaded bands represent the confidence intervals (95%). The higher the participants' scores of Familiarity, the higher the Easiness to Start a Conversation scores.

To summarize, the advantage of the Emotional and Self-Sociality concepts on ESC dimension was not explained by Familiarity. As shown in Fig. 4B, the advantage of Emotional and Self-Sociality Subclusters is present independently of the Familiarity dimension.

*Abstractness ~ concreteness fully captures the words' distribution as a function of new pragmatic-emotional dimensions*

Although many of the dimensions we targeted in our study are commonly employed in studies on conceptual representations, some of these dimensions, like ON, ESC, and Pleasantness, are new. For this reason, we were interested in understanding whether our data would be grouped in new subclusters on the basis of the eight dimensions explored. So, we performed a hierarchical cluster

analysis based on Euclidean distances between our variables. Silhouette's Index (Chouikhi et al., 2015) indicated two as the best number of clusters for our data ( $SI = 0.2$ ). We identified relevant clusters based on their Euclidean distance on the dimensions of ratings, using Ward's clustering algorithm (see Harpaintner et al., 2018; Mazzuca et al., 2020 for similar methods), an agglomerative clustering algorithm that minimizes the total within clusters variance. Cluster 1 included all concrete words but few abstract words, while Cluster 2 included all abstract words but few concrete words. Cluster 1 encompasses words with higher scores on Familiarity ( $M1 = 5.69$ ,  $M2 = 5.35$ ,  $t(214) = 5.10$ ,  $p < 0.0001$ ), CA ( $M1 = 5.71$ ,  $M2 = 4.59$ ,  $t(214) = 23.21$ ,  $p < 0.0001$ ), and WC ( $M1 = 6.44$ ,  $M2 = 6.01$ ,  $t(214) = 19.74$ ,  $p < 0.0001$ ) compared to words in Cluster 2. By contrast, words in Cluster 1 had lower scores on SM ( $M1 = 0.59$ ,  $M2 = 1.17$ ,  $t(214) = 21.525$ ,  $p < 0.0001$ ), ON ( $M1 = 0.72$ ,  $M2 = 1.90$ ,  $t(214) = 33.58$ ,  $p < 0.0001$ ) and ESC ( $M1 = 3.25$ ,  $M2 = 3.72$ ,  $t(214) = 4.93$ ,  $p < 0.0001$ ) compared to words in Cluster 2. The two clusters did not differ for Valence ( $M1 = 4$ ,  $M2 = 4.13$ ,  $t(214) = 0.76$ ,  $p = 0.45$ ) and Pleasantness ( $M1 = 3.69$ ,  $M2 = 3.85$ ,  $t(214) = 1.14$ ,  $p = 0.25$ ). In conclusion, ACs are polarized in a cluster with the following features: high SM, high ON, and high ESC, while CCs are polarized in a cluster with the following features: high WC, high CA, and high Familiarity.

### **Interim conclusions**

In summary, the overall analysis showed that the more pleasant, familiar, and positively valenced were rated the concepts, and the easier to start a conversation with was evaluated. At lower values of the Familiarity dimension, ACs were rated as higher than CCs in the ESC dimension. Moreover, ACs were rated higher on SM, ON, and lower on WC and CA. From the PCA analysis, we extracted two dimensions: one emotional–experiential, one semantic–metacognitive. The first one is composed of Familiarity, Valence, and Pleasantness and stronger predicts the ESC dimension, the second one is composed of WC, CA, SM, ON, and markedly distinguishes AC and CC. The subclusters of ACs with the highest rating in the ESC dimension were the Self-Sociality and emotional subclusters; this result was not explained by the Familiarity dimension. Finally, on the basis of the explored dimensions (Pleasantness, Familiarity, Valence, SM, WC, CA, and ON), we identified through the hierarchical cluster analysis two clusters of concepts. One cluster corresponded to ACs with high SM, high ON, and high ESC, while the other cluster corresponding to CCs with high WC, high CA, and high Familiarity.

## Study 2

Although not included in the preregistration, we performed an additional rating study on the ESC dimension, together with Pleasantness and Valence dimensions targeting an English native sample. Our aim was to verify whether the advantage of AC vs. CC on ESC holds across cultures.

### Method

#### Participants

A total of 40 participants took part in the study in a window between December, 10th 2021 and May, 30th 2022. Participants were recruited via anonymous links either by posting the surveys on social networks (Facebook, Twitter) or spreading the questionnaires through the research team's extended network of acquaintances. English was the native language of all the participants, who were naïve as to the experiment's purpose. Their participation was voluntary and not reimbursed. We decided to include a sample of 20 participants per each survey to recruit a number of participants comparable to Study 1.

#### Materials and procedure

To prevent participants from fatiguing effects, we split the sample of words into two different surveys. Each survey was composed of 40 different words (20 abstract and 20 concrete) to be rated across the three following dimensions: Valence, ESC, and Pleasantness. The two surveys had the same structure, but the 40 words to be rated were different. Since the surveys to be compiled were two, a sample of 40 participants was needed, 20 for each survey. A total of 74 participants were originally contacted, but 34 of them were excluded because they did not complete the questionnaires, likely because of their length (34 participants, 45.95% of the original sample). Of the final sample of 40 participants, 20 participants (15 females and 5 males; M age = 30.8 years, SD = 12.89) performed the ratings of the first 40 words; the remaining 20 (13 females and 7 males; M age = 34.75 years, SD = 15.94) performed the second ratings on the other 40 different words. The study was conducted in accordance with the Declaration of Helsinki guidelines. Ethics permission was granted by the Ethics Committee of the Psychology Department at Sapienza University of Rome (Prot. n. 0001040– 16/11/2020). Participants were informed of the general purpose of the study. Each survey provided informed consent. The study was implemented on Psytoolkit (Stoet, 2010, 2017), in two independent surveys to avoid fatigue effects. Each one presented 20 abstract

and 20 concrete words that participants were required to rate. ACs selected from Villani et al.'s (2019) database were counterbalanced for the four subclusters identified in the database and randomly assigned to one of the two survey ratings. Accordingly, 10 physical, spatio-temporal, and quantitative concepts; 10 philosophical and spiritual concepts; 10 concepts related to the self and social concepts; and 10 emotional and inner state concepts were selected. The 40 CCs were selected from Montefinese et al.'s (2014) database. The surveys required for each concept a rating score for the three dimensions of interest: Valence, Pleasantness, and ESC.

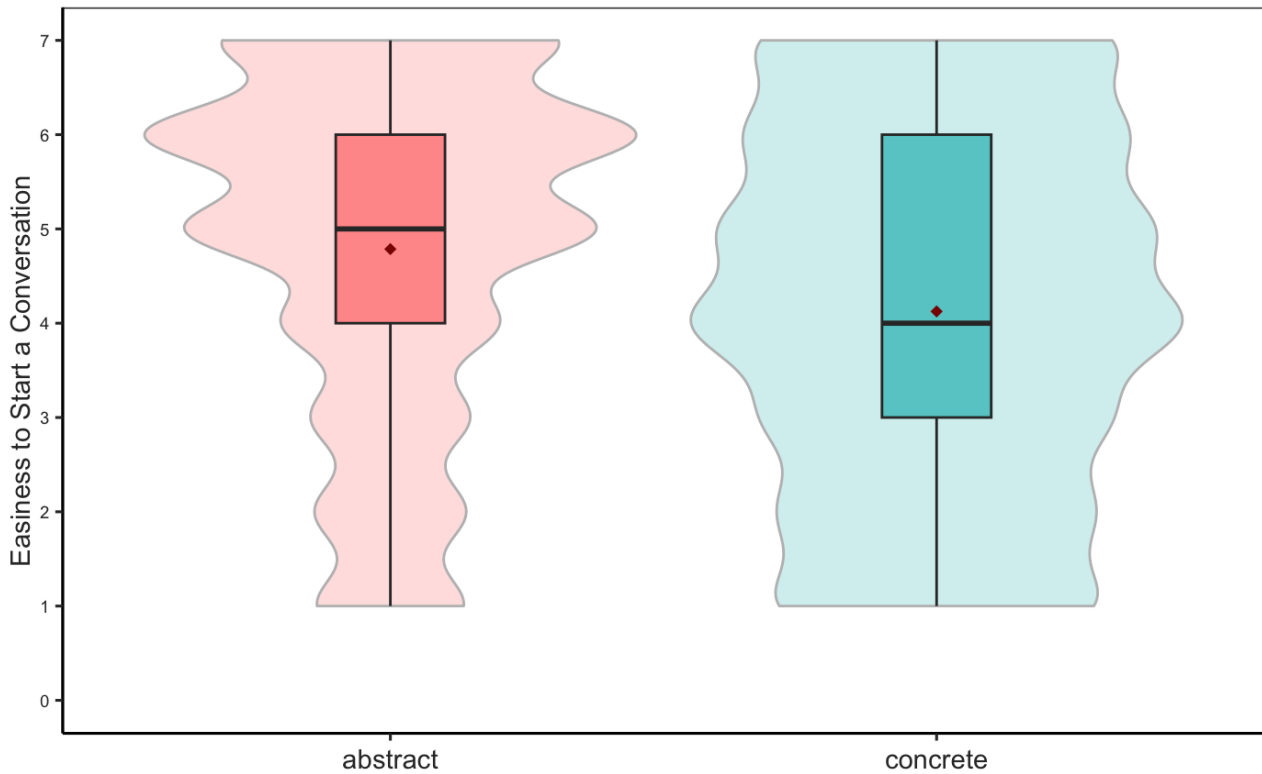
## **Data analysis**

The procedure was the same as Study 1. All data, scripts, and analyses are available at <https://osf.io/tq3nb>.

## **Results**

### *The advantage of AC vs. CC on ESC holds in English*

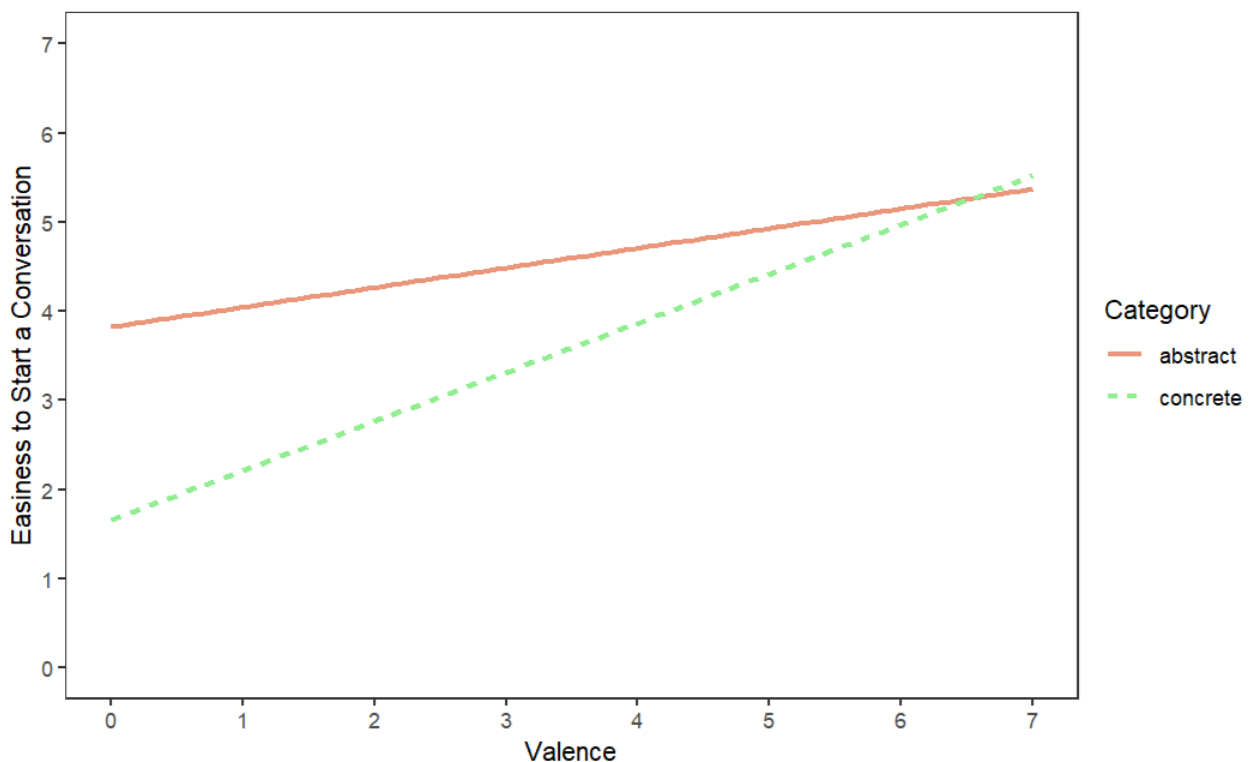
In order to verify whether the advantage of AC vs. CC on the ESC dimension holds in an English sample, we performed a linear mixed-effects model. The model ( $R^2C = 0.46$ ) included ESC as dependent variable, Category of Concepts (abstract, concrete) as fixed factor, and participants and words as random intercepts. The model showed a main effect of Category of Concepts  $F(1,214) = 14.657$ ,  $p < 0.0001$ , with ACs ( $M = 4.79$ ,  $SE = 0.205$ ) rated significantly higher on ESC as compared with CCs ( $M = 4.12$ ,  $SE = 0.205$ ,  $t(77) = 4.141$ ,  $p < 0.0001$ ), (see Figure S4).



**Figure S4.** The graph shows the main effect of Category of Concepts yielded by the model with as dependent variable Easiness to Start a Conversation -ESC-, as fixed effect the Category of Concepts (Abstract, Concrete) and as random intercepts, participants and words. ACs are rated significantly higher as compared to CCs in the Easiness to Start a Conversation-ESC-. Horizontal lines in the boxes indicate the median, upper and lower borders indicate first and third quartiles; red rhombus represents the average of observations, ‘whiskers’ extend to the farthest points that are not outliers.

Then, as for the Italian sample, we investigated whether the effect was modulated by the following dimensions: Pleasantness, Valence, and Familiarity. To avoid multicollinearity, we computed correlations among the continuous predictors to be entered into the model. Valence and Pleasantness were highly correlated ( $R = 0.75$ ,  $p < 0.0001$ ), moreover, Familiarity was correlated with Valence ( $R = 0.14$ ,  $p < 0.0001$ ) and with Pleasantness ( $R = 0.17$ ,  $p < 0.0001$ ). For this reason, we performed three different models with ESC as dependent variable, Category of Concepts as fixed effect, and random intercepts for words and participants, varying for the continuous predictors. In the first one, we entered Valence as continuous predictor (M1); in the second model we entered Pleasantness as a continuous predictor (M2); finally in the third model we entered Familiarity as continuous predictor (M3).

The model ( $R^2C = 0.46$ ) with Valence as a continuous predictor yielded a main effect of the Category of Concepts  $F(1,1529.04) = 32.022$ ,  $p < 0.0001$ , and Valence  $F(1,1183.54) = 71.155$ ,  $p < 0.0001$ . The main effects of Category and Valence were better explained by the two-way interaction Category x Valence  $F(1,1248.31) = 14.599$ ,  $p < 0.0001$ . Simple slope analysis revealed that the slopes of AC (LCI 0.0671 – UCI 0.223) and CC (LCI 0.2821 – UCI 0.461) were significantly different from zero as a function of the factor Valence. The pairwise difference between the simple slopes of AC and CC as a function of Valence was significant (estimate = 0.227, SE = 0.0596,  $t(1249) = 3.807$ ,  $p < 0.0001$ ). The higher the ratings of concepts in Valence, the higher the ratings in the Easiness of Starting a Conversation dimension, the less the difference between AC on CC (see Figure S5).

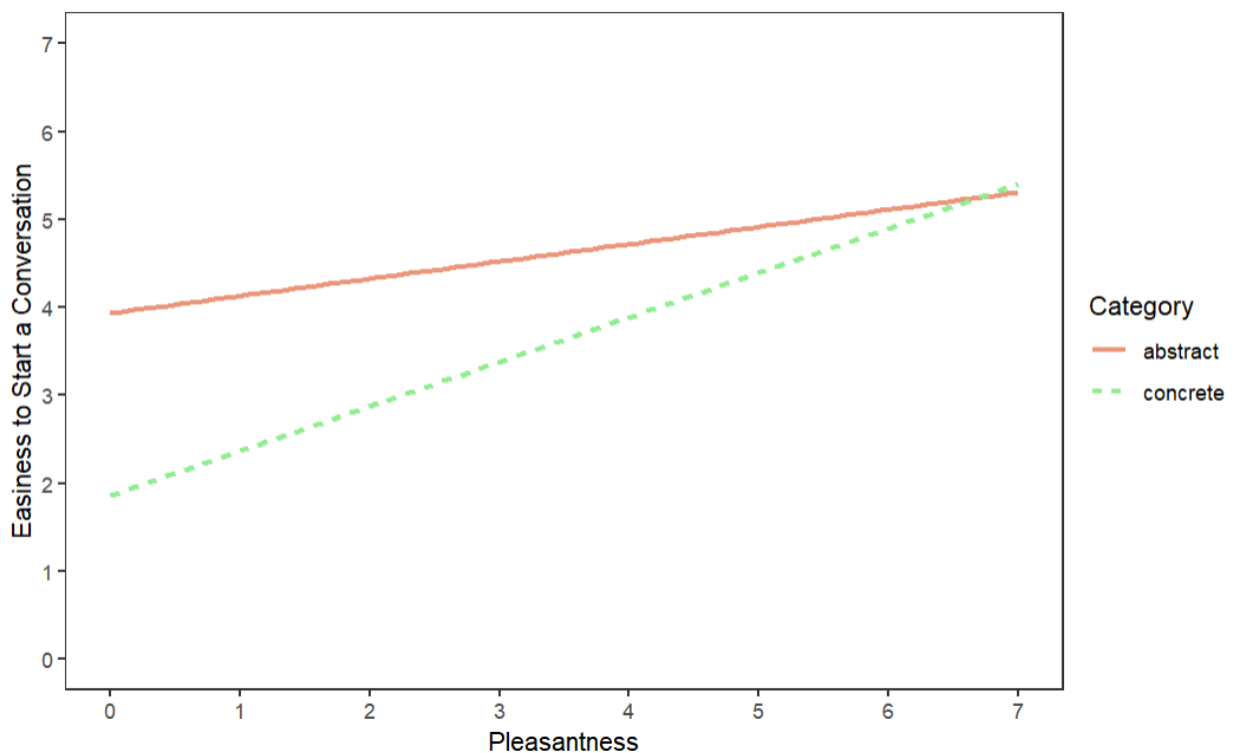


**Figure S5.** The graph shows the predicted values of the outcome variables. Shaded bands represent the confidence intervals (95%). The higher participants' scores of Valence, the higher the Easiness to Start a Conversation scores.

## M2

The model ( $R^2C = 0.49$ ) with Pleasantness as a continuous predictor yielded a main effect of Category of Concepts  $F(1,453.86) = 37.845$ ,  $p < 0.0001$  and Pleasantness  $F(1,1040.54) = 54.160$ ,  $p < 0.0001$ . The main effects were better explained by the twoway interaction Category x

Pleasantness  $F(1,1130.75) = 18.366$ ,  $p < 0.0001$ . Simple slope analysis revealed that the slopes of AC (LCI 0.0158 – UCI 0.167) and CC (LCI 0.2479 – UCI 0.411) were significantly different from zero as a function of the factor Pleasantness. The pairwise difference between the simple slopes of AC and CC as a function of Pleasantness was significant (estimate = 0.238, SE = 0.055,  $t(1137) = 4.269$ ,  $p < 0.0001$ ). The higher the ratings of concepts in Pleasantness, the higher the ratings in the Easiness of Starting a Conversation dimension, the less the difference between AC on CC (see Figure S6).

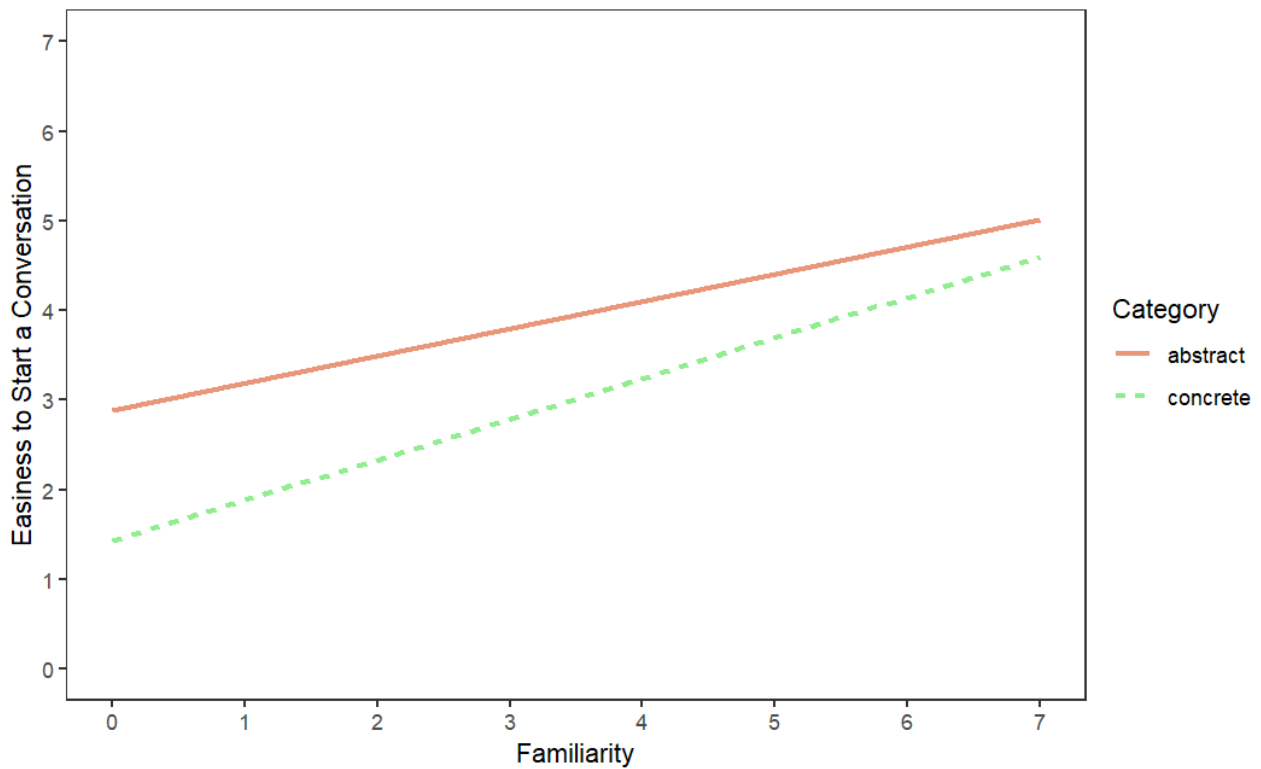


**Figure S6.** The graph shows the predicted values of the outcome variables. Shaded bands represent the confidence intervals (95%). The higher participants' scores of Pleasantness, the higher the Easiness to Start a Conversation scores.

### M3

The model ( $R^2C = 0.49$ ) with Familiarity as a continuous predictor yielded a main effect of Category of Concepts  $F(1,1150.4) = 12.3699$ ,  $p = 0.0004$  and Familiarity  $F(1,1589.2) = 110.9192$ ,  $p < 0.0001$ . The main effects were better explained by the two-way interaction Category Familiarity  $F(1,1564.8) = 4.8518$ ,  $p = 0.0277$ . Simple slope analysis revealed that the slopes of AC (LCI 0.189 – UCI 0.379) and CC (LCI 0.334 – UCI 0.480) were significantly different from zero as a function of the factor Familiarity. The pairwise difference between the simple slopes of AC and CC as a function of Familiarity was significant (estimate = 0.123, SE = 0.0558,  $t(1565) = 2.200$ ,  $p =$

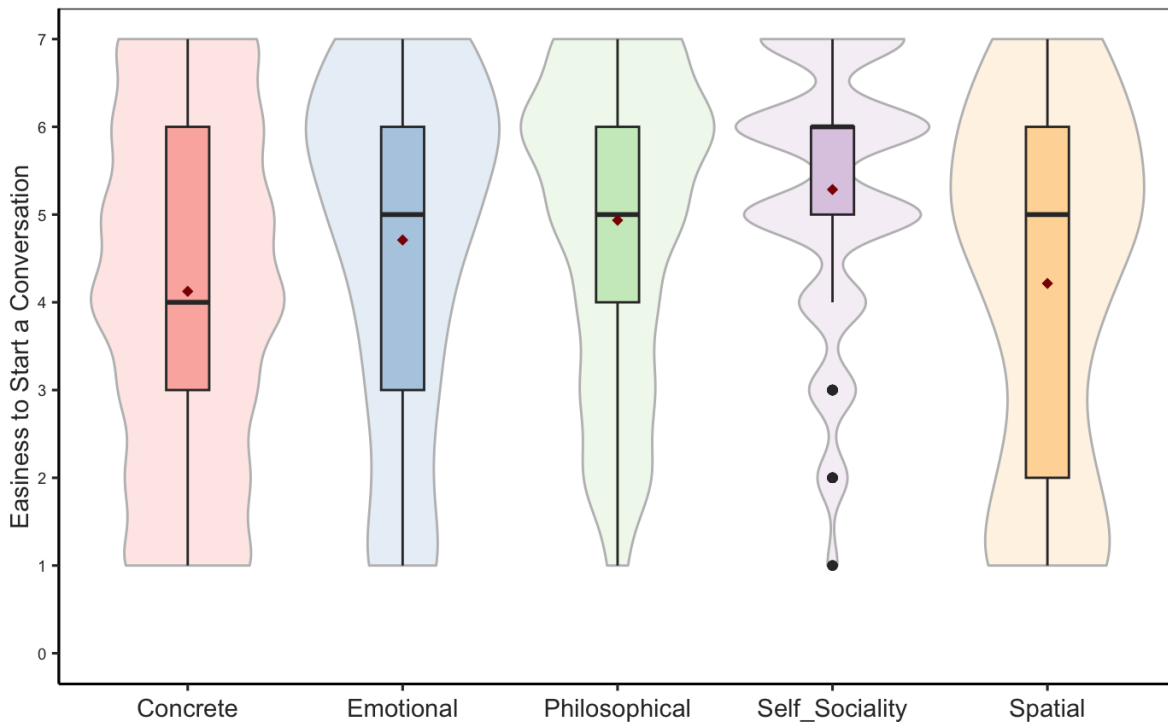
0.028). The higher the ratings of concepts in Pleasantness, the higher the ratings in the Easiness of Starting a Conversation dimension, the less the difference between AC on CC (see Figure S7). In conclusion, the emotional–experiential dimensions positively predicted the ESC but did not explain the advantage of ACs over CCs.



**Figure S7.** The graph shows the predicted values of the outcome variables. Shaded bands represent the confidence intervals (95%). The higher participants' scores of Familiarity, the higher the Easiness to Start a Conversation scores.

*Self-Sociality subcluster is rated higher compared to all the other subclusters on the ESC dimension also in an English sample*

We also explored in the English sample whether specific ACs' semantic subclusters identified by Villani et al. (2019) differ on the ESC dimension. So, we performed a model ( $R^2C = 0.47$ ) with ESC as dependent variable, Subclusters of Concepts as a fixed factor, and participants and words as random intercepts. The model showed a main effect of Subcluster  $F(4,78.11) = 5.768, p < 0.0004$ . Tukey post hoc comparisons indicated that Self-Sociality concepts ( $M = 5.16, SE = 0.219$ ) were rated significantly higher in ESC as compared with CCs ( $M = 4.12, SE = 0.203, t(79.7) = 3.873, p = 0.002$ ). As in the Italian sample, the subcluster which seems to favor starting a conversation is the Self-Sociality (see Figure S8).



**Figure S8.** The graph shows the main effect of Subcluster of Concepts yielded by the model with as dependent variable Easiness to Start a Conversation -ESC-, as fixed factor the Subcluster of Concepts and as random intercepts, participants and words. Self-sociality ( $M=5.16$ ,  $SE=0.219$ ), concepts scored significantly higher on Easiness to Start a Conversation as compared to Concrete concepts ( $M=4.12$ ,  $SE=0.203$ ).

### 3.1.2. Discussion

In the current study, we investigated whether starting a conversation with an abstract rather than a CC is perceived as easier. To this aim, we asked participants to rate AC and CC, on the ESC dimension, along with other classical and less-explored conceptual dimensions. We predicted that ESC might be related to how much the meaning of a word can be negotiated/ discussed –(ON); how much we need the others’ contribution to validate the meaning of a word (SM, Villani et al., 2019)]; how much we feel confident in mastering the meaning of the word (WC), and how much it is easy to think of a context for each word (CA, Schwanenflugel et al., 1988). Results show that regardless of being abstract or concrete, the more a concept was rated as pleasant, positively valenced, and familiar, the easier it was to start a conversation with it. More importantly, participants judged it easier to start a conversation with an AC compared to a CC. The lower the Familiarity, the higher the advantage of AC compared to CC in the ESC dimension. We also found that ACs scored higher on SM and ON, and lower on CA and WC. Indeed, we might need the others’ contribution more to master ACs (Fini et al., 2021; Villani et al., 2019) and they might trigger more discussion and

debates (Mazzuca & Santarelli, 2022) compared to concrete ones. On the other hand, we feel more confident in mastering CCs (Mazzuca et al., 2022) because they are easier to understand and offer a clear, imaginable context/situation as referent (Schwanenflugel et al., 1988) compared to ACs.

Overall, the ESC dimension seems better explained by emotional–experiential components resulting from the combination of Familiarity, Pleasantness, and Valence, extracted with the PCA. Regardless of concepts being abstract or concrete, the higher the Familiarity and Pleasantness, the more positive the Valence, the higher the ESC. When looking at the difference between AC and CC on the ESC dimension, the lower the emotional–experiential ratings, the higher the advantage of AC over CC. The advantage of ACs over CCs in the ESC dimension was also confirmed by a hierarchical cluster analysis, in which two clusters of concepts were identified. The first included mostly ACs with high scores on SM and ON, and the second included mostly concrete words with high scores on CA, WC, and Familiarity. These two clusters did not differ for Pleasantness and Valence. By investigating which kind of semantic subclusters most contributed to the advantage of ACs in the ESC dimension, it emerged that Self-Sociality and emotional concepts were those evaluated as easier to start a conversation with. Crucially, their predominance was not explained by Familiarity. The advantage of AC on CC in the ESC dimension was also replicated in an English sample with another set of stimuli. The subclusters that scored higher in the ESC were again Self-Sociality and the Emotional concepts. The Self-Sociality subcluster significantly differed from the concrete subcluster, while the emotional one showed a tendency in the same direction.

### *ACs “afford” conversations*

The role of conceptual semantics in how we use the concepts during real social exchanges has been poorly investigated. Here, through a rating study, we found that ACs are perceived as better “affording” conversations. The results obtained with an Italian sample generalize to an English-speaking sample. Indeed, ACs were evaluated as easier to start a conversation with than concrete ones. Why might ACs facilitate conversations?

ACs seem to evoke linguistic experiences to a larger extent, as shown by evidence on activation of the mouth motor system during ACs acquisition and processing. Furthermore, they evoke more uncertainty, as revealed by judgments of lower confidence (Mazzuca et al., 2022) and by a higher number of uncertainty-related expressions (Villani et al., 2022). Conversations might be not only the prerequisite for ACs’ acquisition but also the optimal scaffolding to discuss, negotiate, and master complex meanings and to find an intellectual validation.

In previous work, we argued that ACs elicit prosocial behavior due to their indeterminate character (Borghi, 2022; Fini & Borghi, 2019). In this rating study, we consistently found that people feel less confident in knowing their meaning, tend to rely more on others to understand their meaning, and believe their meaning is more debatable and open. This might explain why people tend to think it is easier to start a conversation with ACs. However, our results suggest that such an explanation is not complete. Indeed, we found in both the Italian and the English sample that, among ACs, the kinds of concepts with which it is easier to start a conversation are Self-Sociality concepts (i.e., “politeness,” “art,” “mood”), the ACs which are more related to social matters, and that are more “embodied” than other ACs (Villani et al., 2019).

One can speculate that two alternative explanations are at play. The first leads to starting a conversation from matters more related to sociality. Others would therefore be perceived as people with whom to share experiences. This explains both the higher scores in ESC of AC over CC and the advantage of Self-Sociality concepts over the other ACs. Another, and possibly concurrent, mechanism is more linked to the intellectual role others might play. This mechanism might contribute to explaining the advantage of AC over CC in starting a conversation. We might engage in conversations and dialogs with others when we need their intellectual support to understand better complex concepts, including the more abstract and less-embodied ACs.

Recent results (Villani et al., 2022) of a study in which participants had to respond to sentences simulating a conversation with an acquaintance indicated that Philosophical and Spiritual concepts and Physical Space Time and Quantity concepts seemed to contribute to fuel the conversation over time. Participants were keen to continue the conversation, asking for more turns when the conversation was about Philosophical and Spiritual concepts and Physical Space Time and Quantity concepts. Based on this evidence, we are akin to think that the previously described mechanisms might pertain the relation between ACs’ subclusters (see Villani et al., 2019) and conversational dynamics in different ways. Self-Sociality concepts might favor starting a dialog, by pointing directly to emotional and psychological states, which are part of conventional wisdom and by avoiding opportunities for conflicts and polarizations. Philosophical and Spiritual concepts and Physical Space Time and Quantity concepts might, instead, lead participants to extend the dialog, thanks to intense intellectual exchange.

The emotional–experiential component (Familiarity, Valence, Pleasantness) extracted with the PCA impacts the ESC dimension. The more a concept is familiar, positively valenced, and pleasant, the higher the ESC. Importantly, although such a component seems to play a crucial role in explaining overall the ESC dimension, it does not explain the advantage of AC over CC in the

same. As attested by the hierarchical cluster analysis, ACs are included in a cluster characterized by high scores on SM, ON, and ESC, while CCs are included in a cluster characterized by high scores on WC, CA, and Familiarity. Crucially, regardless of Familiarity, Valence, and Pleasantness, ACs are rated higher in the ESC than CCs.

Finally, three aspects deserve attention. The first one is that the stimuli selection was performed by avoiding too negatively valenced words; this implies that by exploiting a full range of emotional words, the emotional–experiential dimensions might override the advantage of ACs over CCs. The second is that future behavioral studies need to explore the results further. Operationalizing the same experimental question in a lab setting will clarify whether the advantage of ACs over CCs in prompting a dialog also occurs during a real conversation. Moreover, we asked participants to rate “decontextualized” words and that we still do not know whether a number of factors can impact conversational dynamics. For example, different environmental features like affordance richness, outdoor vs. indoor spaces, and personality traits might predict the words we choose to start a verbal exchange with others.

## **Conclusion**

ACs (specifically Self-Sociality) are evaluated to ease the starting of a conversation compared to CCs. This effect is not explained by emotional–experiential dimensions, that is, Familiarity, Pleasantness, and Valence. We suggest that it is instead well captured by the social grounding of ACs (see WAT theory, Borghi et al., 2017, 2018, 2019), which might extend to their use and, more generally, to the pragmatic of communication. Gathering knowledge about new conceptual pragmatic dimensions might be particularly useful for an adequate linguistic stimuli selection in more ecological interactive paradigms to study conversational dynamics.

## **3.2. Digital Connection, Real Bonding: Brief Online Chats Boost Interpersonal Closeness Regardless of the Conversational Topic**

### **3.2.1. Introduction**

As of April 2023, 5.8 billion people (64.6 % of the world population) are estimated as Internet users (Internet and social media users in the world 2023 | Statista). The rapid diffusion of virtual networking powerfully triggered a social change (Wellman et al., 2003). Social networks such as Facebook, Instagram, Whatsapp, Telegram, Twitter, are new platforms for self-expression and allow people to be constantly available to communicate with others (Vorderer & Kohring, 2013). The technological intrusion into the interpersonal sphere caused important psychological and sociological phenomena. More than others, new interactive styles are a direct consequence of the burgeoning of computer-mediated communication (CMC, see Tidwell & Walther, 2002). The digital space allows people to start a conversation about a topic, and create personal relationships. What are the main differences between CMC social interaction as compared to the traditional face-to-face ones? Corporeal proximity is no longer needed for CMCs to take place. When chatting by writing messages on a computer screen, the richness of bodily language cannot contribute to smooth conversational dynamics. All kinds of interpersonal nonverbal signals and paralinguistic aspects (i.e., gaze direction, voice tone, proximity, face expression) typical of face-to-face interactions (see Bavelas, & Gerwing, 2011; Bavelas, Gerwing & Healing, 2014; Holler et al., 2022; Bavelas 2022) are indeed missing (Daft and Lengel 1984; Sproull and Kiesler 1986; Shokley, Richardson & Dale, 2009). Research, however, reveals that CMC allows the maintenance of social ties and the establishment of new relationships among individuals (McKenna, Green, & Gleason, 2002; Parks & Floyd, 1996; Ellison et al., 2014; Grieve et al., 2013; Pettegrew & Day, 2015; Ryan et al., 2017; Wei & Lo, 2006).

In written synchronous CMC, where users engage in real-time conversations by sending messages that appear on the receiver's computer screen, the lack of embodied social feedback is partially compensated by alternative strategies (e.g., repeating the letters of some words, and using emoticons, contractions, and acronyms; Zaferanieh, 2012, see also Walther & Tidwell, 1995;

Walther & D'Addario, 2001). The sequence of the interactive events described above reinforces what we already know about dialogues: regardless of the (un)mediated modalities through which they are realized, conversations are a form of joint activity (Clark, 1996; Garrod & Pickering, 2004; 2021). During a dialogue, two interlocutors try to accomplish the common goal of building new knowledge based on reciprocal conversational contributions. Without collaborating, a conversation becomes fragmented, chaotic, and ambiguous and the information exchanged will not be fully understood by the interlocutors, as they lack a shared interpretative background which can lead both to agreements or disagreements. Even though both face-to-face conversations and written synchronous CMC, are the product of collaboration efforts, the adopted interacting styles are quite different (Tidwell & Walther, 2002). Various theoretical views attempt to frame these differences.

The hyperpersonal perspective of CMC claims that the absence of nonverbal cues, editing capabilities, identity cues, and temporal characteristics may lead interlocutors to favour self-presentation and partner idealization, fostering exchanges that are more intimate than those of face-to-face counterparts. Similarly, the Social identity and Deindividuation Theory (Spears & Lea, 1992) sustains that, in synchronous CMC, reciprocal impressions are formed based on the social categories attributed to interlocutors rather than on interpersonal cues. In line with the Uncertainty reduction theory — (Berger & Calabrese, 1975), Tidwell and Walther (2002) found that, in written synchronous CMC, interlocutors employ more direct interactive uncertainty reduction strategies, such as asking intimate questions and sharing personal information, as compared with face-to-face conversations. Technology-mediated self-disclosure is at least as frequent and as relevant as face-to-face self-disclosure (Nguyen, Bin, & Campbell, 2012) and may lead to increased intimacy (Jiang, Bazarova & Hancock, 2011). In conclusion, although qualitatively different from face-to-face conversations, written synchronous CMC seems to foster interpersonal communication. A meta-analysis on the relation between digital communication and psychological well-being drew the conclusion that synchronous conversations —instant messaging (IM)— lead to experiencing a positive sense of connection (Liu et al., 2019).

How do people experience synchronous CMC? Is it enough to talk for a few minutes with strangers to create a sense of closeness with them? To date, no research has addressed the interpersonal dynamics promoted by a short written synchronous CMC. As we already pointed out, since a conversation is a form of joint action, it might be sensitive to the same rules orchestrating joint actions. A recent study revealed that participants preferred to maintain a shorter distance from an (unfamiliar) individual when they imagined acting jointly (e.g., cutting together the same branch of a tree) compared to when they imagined acting in parallel (e.g. cutting separate branches of a

tree) (Schmitz & Reader, 2023). Building on this evidence, we might expect that the same pattern of results holds also when the joint action consists in a written synchronous CMC. We hypothesize that two interlocutors might experience a sense of physical-interpersonal closeness after having conversed about a specific topic for a few minutes, as compared with the condition in which two people are simply aware of their reciprocal presence on the social platform, without interacting.

Moreover, Lahnakoski et al. (2020) found that the stronger predictor of the quality of interaction is the interpersonal distance between the interactors, with increased distance associated with lower enjoyment. Thus, we reasoned that variables like perceived pleasantness, difficulty attributed to the dialogue, intimacy, commitment, and self and other contributions to the conversation might all affect the quality of the interaction, and consequently the interpersonal distance among the interlocutors. As the body is not the privileged channel through which contents are conveyed in written synchronous CMC, the quality of the interaction might depend on the topic of the conversation and the pragmatic digital rules governing the exchange. Psychological interpersonal closeness or how much Self and Other identities representations have in common might capture the psychological closeness triggered by synchronous CMC. This is often measured using the Inclusion of Other in Self — iOS— scale, developed by Aron and Aron (1992,1997). Some studies have already exploited the iOS scale to measure interpersonal closeness after a social interaction (Miller et al., 2019; Morrison et al., 2020). According to Aron and Aron (1997), the integration of what the other shares takes the name of ‘self-expansion’ and corresponds to greater relationship satisfaction and commitment. As we claimed above, in synchronous CMC, social signals in interpersonal linguistic exchanges pass through the conversational content, and its chronometric (Walther & Tidwell, 1995) and typographic cues (Walther & D’Addario, 2001).

In this regard, we hypothesize that the semantics of the conversational topic might somehow serve a pragmatic role: depending on the discussed topic, the interlocutors might be involved in different interpersonal dynamics resulting in a different degree of perceived closeness. Semantic knowledge is stored through concepts. Concrete concepts, (e.g., “cat”, “chair”, “glass”), evoke directly sensorimotor experiences, because they all point to an objectual reference and are easily imaginable (Paivio, 1969). Abstract concepts (e.g., “justice”, “Interactivedom”, “peace”) (Borghi et al., 2017, 2018 a,b, 2019 a,b; Dove 2021, 2022; Pexman et al., 2023), instead, typically lack of an objectual reference, are ubiquitous (Lupyan & Winter, 2018) and are characterized by stronger socio-linguistic components (Barsalou & Wiemer-Hasting, 2005). We can the meaning of abstract concepts only when they share a socio-cultural sense-making background that provides interpretative coordinates. Finally, they evoke more interoception and emotions (Connell, Lynott &

Banks, 2018; Vigliocco et al., 2014; Ponari, Norbury & Vigliocco 2018; Fingerhut & Prinz, 2018). The Words as Social Tool theory (WAT, Borghi et al., 2014, 2017, 2018a,b, 2019a,b, 2022, 2023) proposes that abstract concepts as compared with concrete ones, would be acquired and processed more through social interactions (for theoretical views see, Fini et al., 2019, 2023a; Banks et al., in press), as Fini, Era et al., (2021) tested by applying an interactive paradigm (Era et al., 2023; Boukarras et al., 2021). In addition, a rating study has demonstrated that people perceive it as easier to start a conversation with abstract as compared with concrete topic, supporting the idea that being acquired through linguistic exchanges, abstract concepts would also better afford linguistic interaction (Fini et al., 2023b). Importantly, the WAT theory proposes social metacognition — asking the other's help to complement the knowledge about which people feel uncertainty— as a potential mechanism involved during both the acquisition and the processing of abstract concepts. If this is the case, we would expect an intellectual negotiation, aimed to dispel a higher uncertainty, as well as more interactive exchanges, to happen during conversations about abstract concepts (Villani et al., 2022). We might conceptualize linguistic exchanges during conversation about abstract topics, as coordinated goal-directed and bidirectional actions, where the goal consists of enriching each other with reciprocal knowledge.

In light of this, we can expect that the interpersonal dynamics between the interlocutors might be modulated by the conversational topic: abstract versus concrete. By recognizing the other's contribution, people may commit more to better coordinate with the interactor (Fini & Era, 2021) and this might lead to a feeling of physical and psychological closeness. In the current study, we developed a new interactive paradigm requiring conversations on online platforms. The experiment included an interactive condition, in which participants conversed in dyads, and a noninteractive condition, in which they wrote about concrete or abstract concepts by knowing that on the other side of the screen there was a person doing the same and that, later, they would read what the other had written. In the non-interactive condition, thus, participants knew that they were writing contemporaneously to another person, and on the same content, but did not interact with each other. After every conversation/verbal production, the psychological distance and the physical/interpersonal closeness between the actors were measured. To measure the psychological distance, we used the Inclusion of the Other in the Self (iIOS) Scale (Aron et al., 1992). To measure closeness, we used a Visual Analogue Scale (VAS), where the distance between two avatars embodying the interlocutors could be adjusted and participants were asked to stop their own or the partners' avatar when they felt at ease with the reached interpersonal distance (Lisi et al., 2021). Together with these measures, participants were asked to evaluate on a Visual Analogue Scale

(VAS) the Pleasantness, the Intimacy, the Perceived Difficulty experienced during the verbal production, and the Self contribution, plus the Other contribution and the Commitment in the experimental interactive condition (covariates). We expect an increased physical and psychological closeness in the interactive compared with the non-interactive condition, in addition, we predict that all the measured covariates might intervene to further modulate the effects. Regarding the conversational topic, we hypothesize that starting to talk about abstract compared to concrete concepts, might require more of the other's contribution, and such conversational cooperative attitude might result in an increased physical and interpersonal closeness among the interlocutors.

## **Experiment 1**

### **Participants**

Sixty-eight participants were recruited for the study. Five participants were excluded from the Interactive task as outliers (see 'data analysis' section), so that the final sample included sixty-three participants (25 males, 38 females, group average age =  $25.84 \pm 5.28$  years; group average years of education =  $15.71 \pm 2.8$ ). The sample size was selected through a prospective power analysis performed with the software More Power (Campbell and Thompson, 2012). We inserted as expected effect size a large partial eta squared value (0.14) as observed in Soliman et al., 2015, where it was tested whether interpersonal motor interactions increase self-other overlap. Our analysis indicates that a within-subject design with one two-level factor, a power of 0.90 and a partial eta squared of 0.14 requires a sample size of 68 participants. The experimental protocol was approved by the ethics committee of the Department of Dynamic, Clinical Psychology and Health Studies, Sapienza University of Rome (Prot. n. 0001040 16/11/2020), and was carried out in accordance with the ethical standards of the 1964 Declaration of Helsinki and later amendments. Participants gave their written informed consent to take part in the study and were naïve as to the purposes of the experiment.

### **Experimental task**

#### *Stimuli*

The stimuli consisted of eight abstract (i.e., liberty, logic, challenge, mystery, future, ideal, fate, enigma) and eight concrete concepts (i.e., alarm clock, crown, station, diamond, crystal, canoe, muscle and bookcase). Abstract concepts were selected from the Italian database by Villani and

colleagues (2019); concrete concepts were instead extracted from Della Rosa et al. (2010) Italian database. Abstract and concrete concepts were matched for some linguistic parameters, such as Familiarity (e.g., Barca et al., 2010), based on the rating scores included in Della Rosa et al. (2010) and Villani et al. (2019) databases (Concrete:  $M = 559$ ,  $SD = 101.10$ ; Abstract:  $M = 518$ ,  $SD = 61.71$ ;  $t(14) = 0.97$ ,  $p = 0.22$ ). Valence rating scores were obtained by asking two different groups of 26 and 30 participants to evaluate on a 7-point Likert scale 35 abstract and 35 concrete concepts each in terms of Valence. Evaluations were provided through the compilation of two online surveys implemented on Google Moduli. The sample size for each survey was estimated in line with that of previous works (i.e., a minimum of 20 participants - Falcinelli et al., submitted; Villani et al., 2019). Then, 8 abstract and 8 concrete concepts not differing in their valence (Concrete:  $M = 4.01$ ,  $SD = 0.6$ ; Abstract:  $M = 4.25$ ,  $SD = 0.86$ ;  $t(14) = -0.63$ ,  $p = 0.5$ ) were selected.

### *Online Chat Task*

Pairs of same-gender participants were asked to chat online for five minutes through the ICQ platform (<https://icq.com/desktop/en?>). Participants performed two main tasks, i.e. one interactive and one noninteractive. In the interactive task, each participant was paired with another participant to chat— by writing on the computer screen— about two abstract concepts, and then with a different participant, to chat about two concrete concepts. The order of the chat topic (abstract/concrete concepts, Category Factor) was counterbalanced across pairs. Before starting the experiment, different experimenters took care of calling each participant, to guide him/her in understanding the task procedure. Participants were required to switch off their mobile phone and to close every window on their computer, but the ICQ platform. Moreover, participants were instructed to chat about a specific topic assigned by the experiment, taking the turn at least two times each, and writing at least two lines of text for each turn. The total amount of time given to each pair to chat about a specific topic was 5 minutes. To promote a rich dialogue, participants were told that the fluency and the relevance of their chat would have been evaluated. Participants were not allowed to use emoticons while writing. Once on the ICQ platform, participants were introduced to each other and started to chat.

In the non-interactive task, participants were instructed to write alone on the ICQ platform about a specific topic assigned by the experimenter. As for the interactive task, participants had to write in the chat at least two different sentences of two lines each. The total amount of time given to each participant was 5 minutes. To promote a rich production, participants were told that the fluency and the relevance of their writing would have been evaluated. Participants were not allowed to use

emoticons while writing. Once on the ICQ platform, each participant was introduced to a fake partner (played by the experimenter) and was told that the chat would not allow the person on the other side to read what the other was writing. Indeed, while writing, participants received the feedback “The other is writing...” from the platform page but did not see any text displayed. Thus, participants were writing the text alone, however, they were induced to think that another person was doing the same on the other side and that, at the end of the experiment, they would be allowed to see what the other wrote. Each participant performed both the interactive and non-interactive task, separated by a 2/8 weeks interval and their order was counterbalanced across participants. In both the non- interactive and the interactive task, each participant chatted/wrote about two abstract and two concrete concepts with two different other participants and, once half of the sample was tested, two different abstract and concrete concepts were used as a topic, to increase stimuli variability.

After each chat, participants answered to:

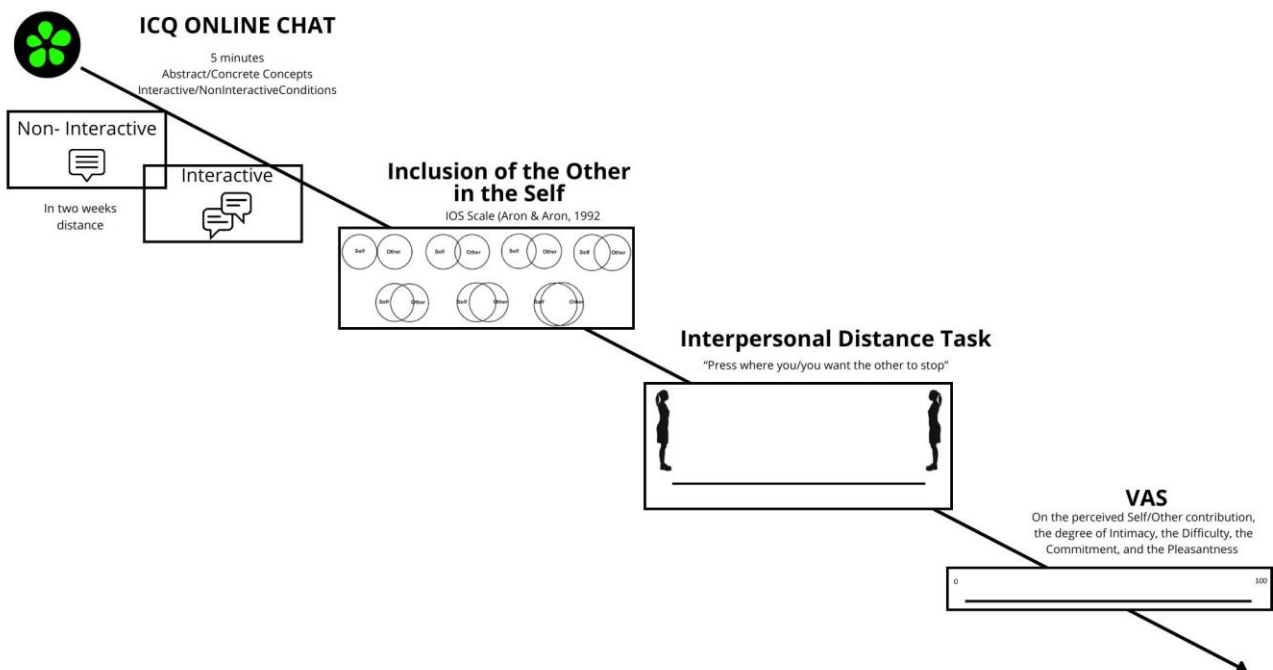
i) An interpersonal distance task (Lisi et al., 2021). The task allows measuring the physical interpersonal closeness among interlocutors. A silhouette representing the participant was selected based on each participant’s gender. The silhouettes were pictured in a standing position on a marker at the left end of a line and facing the right end of the line. The virtual avatars were realized using Make Human Community 1.2.0 (<http://www.makehumancommunity.org>); the pictures of the avatars were taken using Unity v.2019.4.15f1 (<https://unity.com>). Participants were asked to indicate a measure of passive and active distance. In the passive distance measurement, participants were provided with the following instructions: “Imagine that you are the person on the left of the line and that you cannot turn nor move. Then, imagine that the person you just chatted with, depicted on the right side of the line, begins walking toward you. You should indicate how close you would allow this person to approach you while still being comfortable with that distance. To indicate where the other should stop, click on the horizontal line. Then, press ‘Next’ to proceed”. In the active distance

measurement, participants were presented with the following instruction: “In another condition, you will be asked to indicate how close you would get to the other person by imagining that you are moving in the same horizontal line. Click to report where you would like to stop, while still being comfortable with that distance. Then press ‘Next’ to proceed”.

ii) The Inclusion of Other in Self (iOS) Scale (Aron & Aron, 1992), aimed at measuring psychological closeness among interlocutors. In the scale, participants are presented with a single item, where six configurations are displayed. In each configuration, two circles, representing the

Self and the Other, are positioned at an increasing smaller distance, until they overlap. Participants are asked to indicate which configuration best represented their perceived relationship with the person they just chatted with.

iii) Different VAS scales, ranging from 0-100, regarding some aspects of the previous experience of chatting: the perceived Other contribution, the Self contribution (only for the Interactive task), the degree of Intimacy, the Difficulty, the Commitment, and the Pleasantness of the chat.



**Figure 1.** Timeline of the experimental procedure. Participants were asked to chat contemporaneously with another person (interactive condition), or to individually write (non-interactive condition ) about an abstract or concrete topic. After each chat, participants were asked to choose the most representative configuration of the Inclusion of Other in Self (iOS) Scale, to perform the interpersonal distance task and some VAS scales regarding the perceived Other contribution, the Self contribution (only for the Interactive task), the Intimacy, the Difficulty, the Commitment and the Pleasantness of the previous chat.

### *Data availability*

All data and analysis code is available under the following publicly accessible Open Science Framework repository: <https://osf.io/m39br/>

### *Data analysis*

Data were analyzed through R (version 4.1.1 - R-Core Team, 2019) in RStudio (version 4.3.1). ‘Tidyverse’ (version 1.3.2 - Wickham et al., 2019) R’s package was used for data processing and visualization. For creating multi-plot arrays, we used ‘patchwork’ (version 1.1.1 - Pedersen, 2020) and ‘gridExtra’ (version 2.3 - Auguie, 2017) R’s packages. ‘Joy’ or ‘ridge’ graphs were plotted with ‘ggribes’ (version 0.5.3 - Wilke, 2021) R’s package. Finally, Bayesian generalized linear mixed effects models were fitted using ‘brms’ (version 2.16.2 - Bürkner, 2017) R’s package. Before analysing the data, we first assessed whether the closeness ratings differed between the ‘active’ and ‘passive’ questions. A simple Spearman’s rho correlation showed that the two were highly correlated ( $\rho = 0.92$ ), which made us choose to combine the two types of questions into a singular ‘closeness’ variable by taking the average. A total of 8 participants (80 data points, 7.4% of the total data) were excluded from all further analysis, based on the criterion that they were below or above 2 standard deviations from the sampled average in the physical/interpersonal distance.

Our main analysis focuses on two dependent variables: IOS and interpersonal closeness. For all models where the dependent variable is the Inclusion of Other in Self (IOS) Scale, which is ordinal with six scale points, we use a mixed ordinal logistic regression, specifically, cumulative link models (Bürkner & Vuorre, 2019) with logit link function. For all models where the dependent variable is “closeness,” which is a visual analogue scale bounded by [0, 1], we used mixed beta regression with mean/phi parametrization. The beta distribution is commonly used for dependent variables that are continuous but bounded to a finite range, such as proportions and percentages. Our first analyses are targeted at assessing the IOS scale. In a first analysis, we regress all covariates (pleasantness, commitment, intimacy, difficulty, self-contribution, other-contribution) onto IOS in an ordinal logistic regression model with random intercepts for participants and items and by-participant varying slopes for all covariates. This analysis was repeated for the interpersonal closeness measure (mixed beta regression, random effects structure the same). These analyses also informed our choices about which covariates to include in the main model. We then assessed whether closeness and IOS differ as a function of condition (interactive versus non-interactive). For both of these variables, models contain the critical fixed effects “condition” and no other fixed effects. The random effects are: subjects and items, with both random intercepts and random slopes for the condition effect. The main two models were IOS (cumulative ordinal) or interpersonal closeness (beta) regressed onto the fixed effects ‘concept category’ (abstract versus concrete), the continuous covariate ‘other contribution’ (centered), as well as their interaction, ‘category \* other’s contribution.’ We centered the continuous other contribution covariate to more easily interpret the abstract versus concrete category effect (at the mean of IOS). We decided to leave the default

treatment coding system for ‘category’ (0 = abstract, +1 = concrete) as we are not specifically interested in a main effect of the Inclusion of Other in Self scale. Random effects included random intercepts for participants and items, as well as by participant varying slopes for all fixed effects, and by-item varying slopes for category. As discussed in the results section, the fact that abstract concepts also showed a trend to lead to slightly more difficult interactions made us decide to include this as an additional covariate, which is the final model we used. We used default priors assigned by brms, which are reported in the Open Science Framework repository (<https://osf.io/m39br/>).

Importantly, this involves flat priors on beta coefficients, in close analogy to a corresponding frequentist model. All models were fitted with Markov Chain Monte Carlo, with 4 chains 3,000 iterations used for posterior sampling (3,000 warm-up samples were discarded). All models converged adequately (all Rhat = 1.0, all ESS > 1,000). We performed posterior predictive simulations (100 draws) to assess whether our models could plausibly have generated the data, which was the case for almost all models, except for a few models showing minor deviations between the simulated and raw data.

## Results

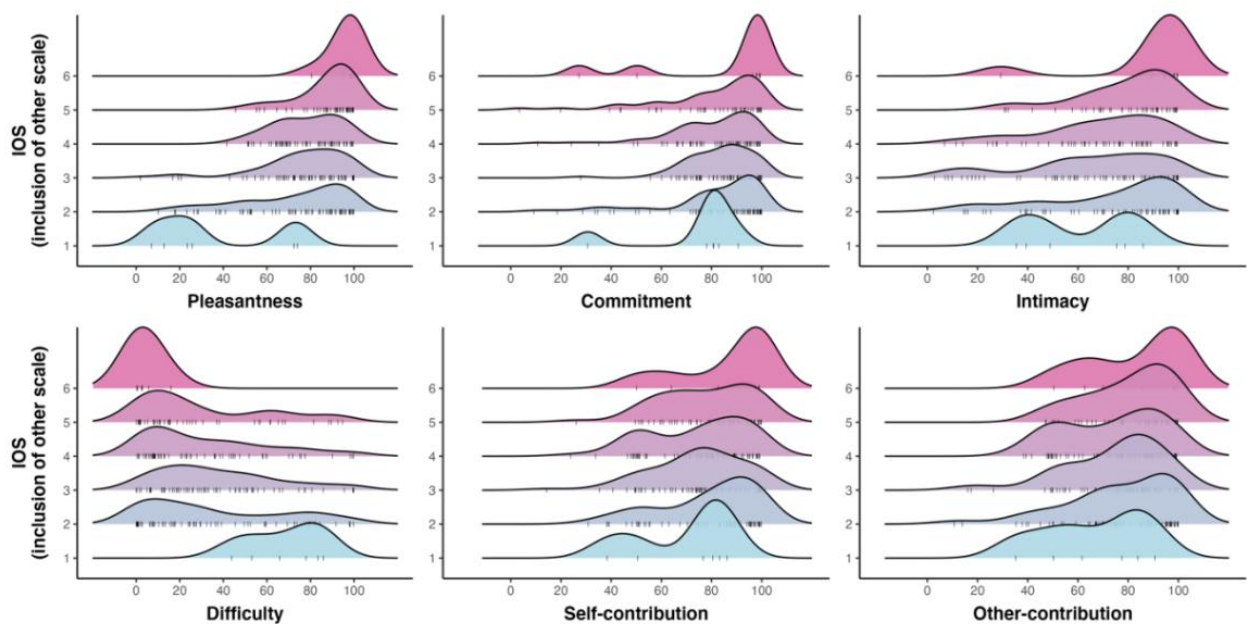
### *Covariates*

Before looking at any condition and concept differences, we will explore how the covariates behave with respect to each other, followed by how all covariates behave with respect to the iOS scale and interpersonal closeness. Table 1 shows Pearson's correlations between all covariates. The highest correlation was between ‘self-contribution’ and ‘other-contribution’ ( $r = .67$ ), followed by ‘self-contribution’ and ‘commitment’ ( $r = .65$ ), and ‘other-contribution’ and ‘pleasantness’ ( $r = .57$ ). The ‘difficulty’ variable was either not correlated with any of the other variables, or negatively correlated, specifically, with ‘pleasantness’ ( $r = -.29$ ), and more weakly so, with ‘intimacy’ ( $r = -.14$ ).

**Table 1.** Pearson’s correlations between all the covariates, i.e., Pleasantness, Commitment, Intimacy, Difficulty, Self-Contribution and Other-Contribution

	Pleasantness	commitment	intimacy	difficulty	self-contribution	other-contribution
pleasantness		.48	.47	-.29	.43	.57
commitment			.45	.00	.65	.47
intimacy				-.14	.47	.41
difficulty					.01	.01
self-contribution						.67
other-contribution						

Figure 2 illustrates the distribution of each covariate for all levels of iOS. The figure shows that pleasantness (top left) differed greatly as a function of iOS. The highest iOS scale point had a pleasantness rating (M = 88.0) that was 79% higher than the lowest point (M = 49.2). For commitment (top middle), the differences were much less pronounced, with the highest iOS scale point having a commitment rating of M = 87.2, compared to only M = 71.0 for the lowest value. Intimacy on the other hand differed more strongly as a function of iOS, and also showed considerably more spread than the other two variables considered so far (top right). The highest iOS scale point had an average intimacy of M = 84.4; the lowest one of M = 59.3).

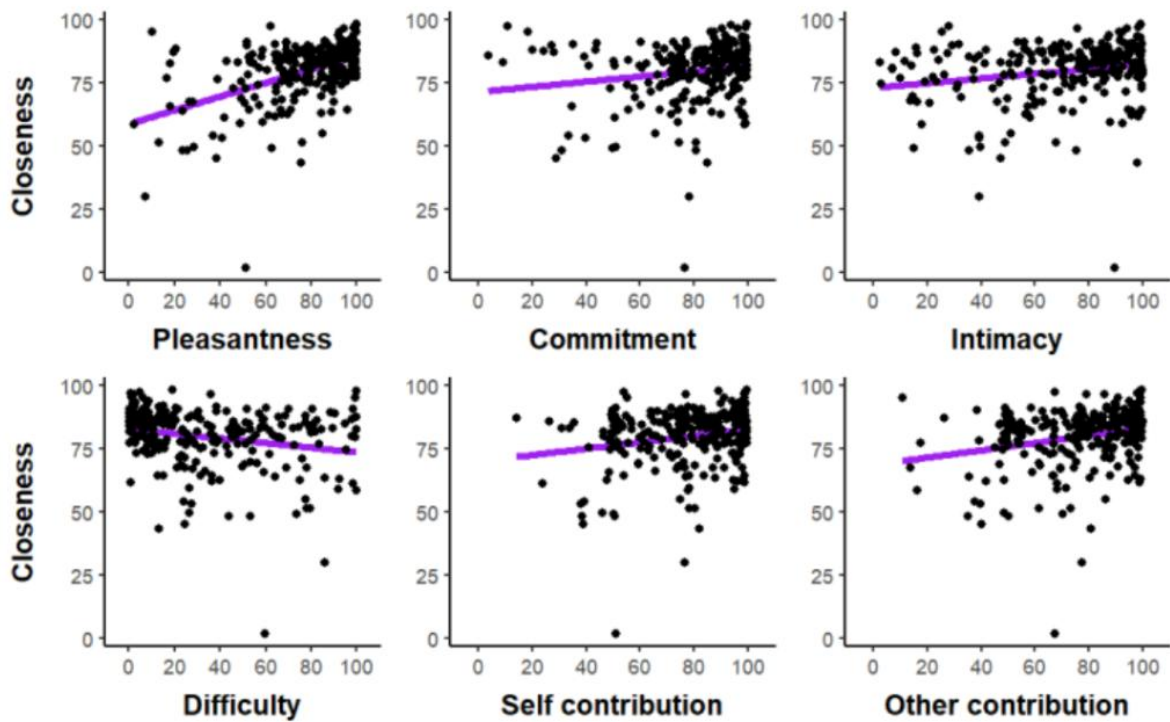


**Figure 2.** ‘Ridge plots’ showing how all six covariates differ for different levels of the Inclusion of Other Scale (IOS).

The highest iOS scale point was associated with the lowest difficulty level ( $M = 3.98$ ), and the lowest iOS scale point was associated with the highest difficulty level ( $M = 68.4$ ). A look at Figure 1 (bottom left) reveals a noteworthy pattern however: difficulty only differed starkly for the first and last iOS scale points, with less stark differences for the intermediate levels (2-5), and considerable spread. Either way, these results show that interactions judged to be difficult lead to lower feelings of other-inclusion. For the highest iOS scale point, self-contribution scores ( $M = 86.1$ ) were somewhat higher than the lowest IOS scale point ( $M = 69.3$ ), although this pattern was not very pronounced (Figure 2, bottom middle). The other-contribution scores (Figure 2, bottom right) were also higher for the highest iOS scale point ( $M = 83.0$ ) than the lowest one ( $M = 66.6$ ), with a similar spread than the self-contribution covariate. To assess uncertainty with respect to the patterns seen in Figure 3, all covariates were regressed onto the iOS scale (with by-participant-varying random slopes for all covariates and word random intercepts). The ordinal logistic regression model revealed that pleasantness most strongly affected IOS (logit coefficient =  $+0.10$ ,  $SE = 0.02$ ), with a 95% credible interval of this coefficient which was far away from zero:  $[0.07, 0.14]$ . The posterior probability of this effect being of the same sign was  $p = 1.0$  (that is, every single posterior sample indicated an effect of the same sign). This means that given this model, data, and priors, there is little uncertainty about pleasantness being positively associated with iOS. The only two other variables for which there was relatively little uncertainty included intimacy and difficulty. Intimacy was positively associated with iOS ( $+0.02$ ,  $SE = 0.01$ , 95% credible interval:  $[0.00, 0.04]$ ), with a posterior probability of  $p = .97$  of the coefficient having the same sign. Difficulty was negatively associated with iOS ( $-0.02$ ,  $SE = 0.01$ ,  $[-0.04, 0.00]$ ), with a posterior probability of  $p = .96$  of being of the same sign. Thus, higher pleasantness and intimacy, but lower difficulty, is associated with higher Inclusion of Other in Self (iOS) values.

We repeated the same analysis for our second dependent variable, interpersonal closeness, which we regressed onto the same covariates (see Figure 3). A similar set of covariates were associated with low uncertainty. Pleasantness was positively associated with interpersonal closeness ( $+0.01$ ,  $SE < 0.01$ ,  $[0.01, 0.02]$ ), with a very high posterior probability of this coefficient being of the same sign,  $p=1$ . Difficulty was negatively associated interpersonal closeness ( $-0.01$ ,  $SE < 0.01$ ,  $[-0.01, 0]$ ), with a very high posterior probability of being negative,  $p= .99$ . In contrast to the iOS model, however, there was no indication of an effect of intimacy on interpersonal closeness, with the

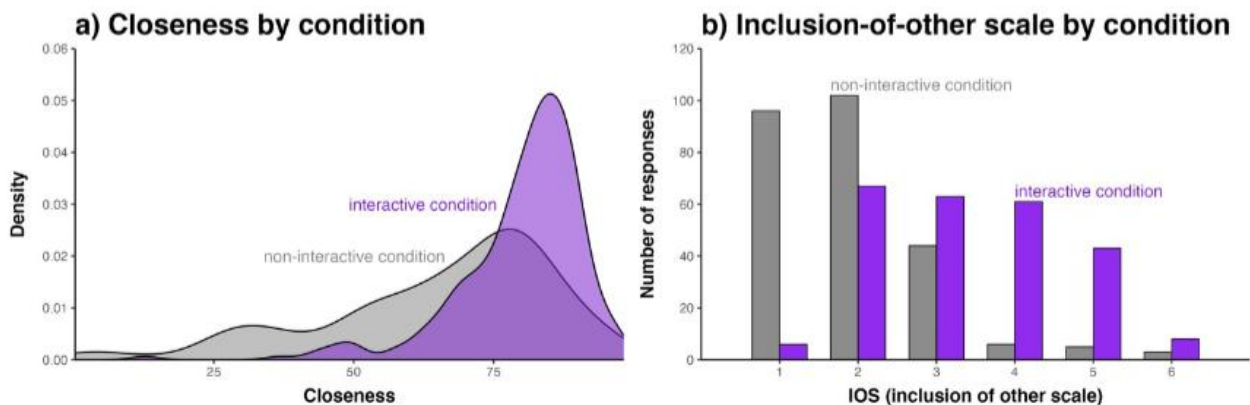
coefficient firmly centered on zero ( $+0.0$ ,  $SE < 0.01$ ,  $[0, 0]$ ), and a posterior probability of the effect being of the same sign that was not informative,  $p = 0.53$ .



**Figure 3.** ‘Regression plots’ showing how all six covariates differ for different levels of interpersonal closeness (VAS).

#### *Effects of interactive versus non-interactive condition*

In this section, we assess how iOS and interpersonal closeness depend on the difference between interactive and the non-interactive condition. Participants felt closer to their partner in the interactive condition ( $M = 80.0$ ) than in the non-interactive condition ( $M = 65.5$ ), as shown in Figure 4, Panel a). The mixed beta regression analysis indicates a positive effect of the interactive condition on interpersonal closeness (logit coefficient =  $+0.72$ ,  $SE = 0.12$ ) for which the 95% credible interval is far away from zero:  $[0.48, 0.97]$ . Given the model and this data, the posterior probability of the interactive condition leading to a positive effect on interpersonal closeness is  $p = 1.0$  (the entire posterior distribution lies above zero, with not a single sample indicating an effect of opposite sign).

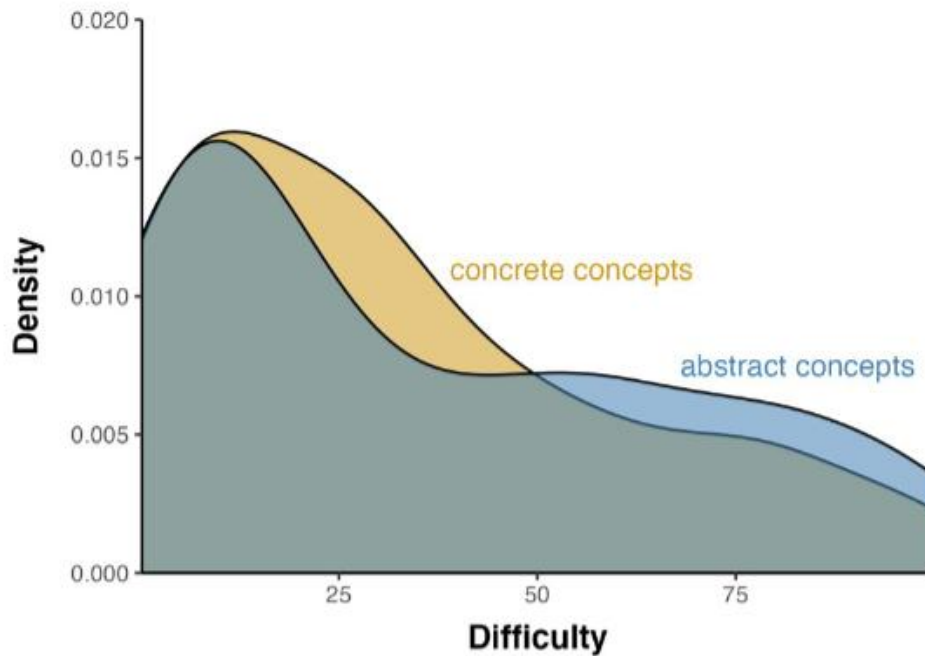


**Figure 4. Panel a)** Distribution of interpersonal closeness values (Visual Analog Scale). **Panel b)** IOS values as a function of condition.

Participants also gave higher iOS score ratings in the interactive condition ( $M = 3.37$ ) than the non-interactive condition ( $M = 1.95$ ) (as shown in Figure 4, Panel b). The mixed ordinal regression analysis indicates a positive effect ( $+4.22$ ,  $SE = 0.48$ ) with a 95% credible interval far away from zero:  $[3.32, 5.18]$ . Given the model and this data, the posterior probability of this coefficient being of the same sign was also  $p = 1.0$

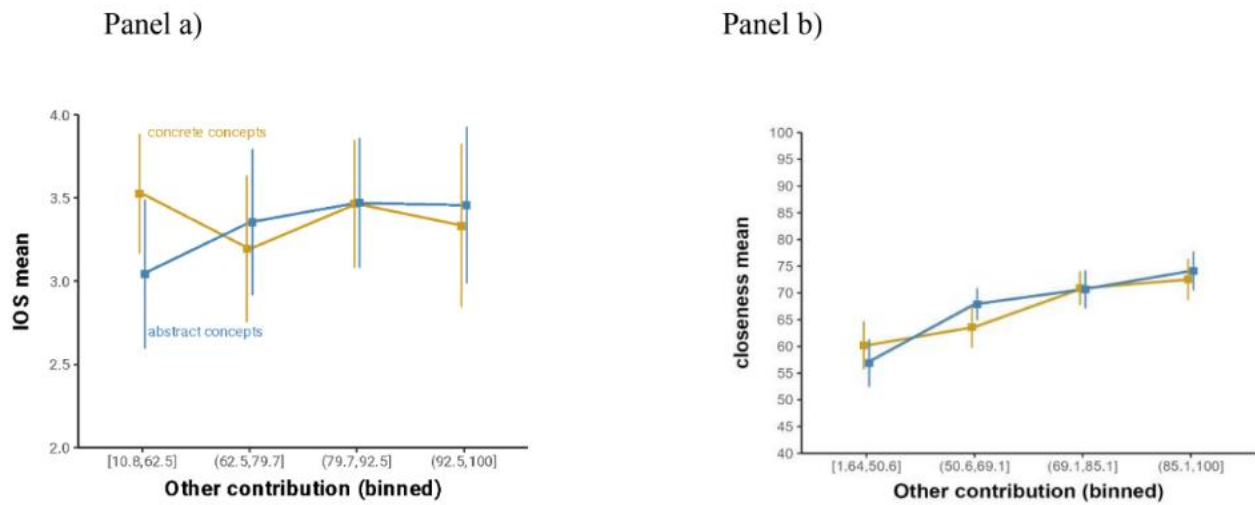
#### *Effects of abstract versus concrete concepts in the interactive condition only*

Before advancing to the main model, we assessed the influence of abstract versus concrete concepts on perceived difficulty. Abstract concepts are perceived to lead to more difficult conversations ( $M = 36.6$ ) than concrete concepts ( $M = 32.6$ ) (see Figure 5). A mixed beta regression analysis on difficulty as the dependent measure indicates a positive effect of abstractness on perceived difficulty (logit coefficient =  $+0.07$ ,  $SE = 0.23$ ), with a 95% credible interval that is clearly lopsided towards higher values  $[-0.34, +0.47]$ , but that also clearly includes zero. The posterior probability of abstract concepts leading to higher difficulty ratings is  $p = 0.63$ . This means that although there is a numerical trend, given the model and the data, it is quite plausible that abstract concepts could also not increase perceived difficulty. Nonetheless, due to the numerical trend and the concern that people may only feel the need to depend on others because abstract concepts are more difficult, we decided to include difficulty as a covariate to see whether abstract concepts lead to higher iOS (or perceived closeness) even when controlling for difficulty. It should be stated, however, that the results reported below hold regardless of whether difficulty is or is not included as a covariate.



**Figure 5.** Distribution of difficulty ratings between concrete and abstract concepts.

Next, we assess our hypothesis, which is that there is an interaction between abstract concepts and the other contribution scale with respect to iOS score. Figure 6 displays this complex relationship via binning the other contribution variable into four equal-sized groups, with separate averages for concrete and abstract concepts. The standard errors are simple standard errors of the mean, which only serve heuristic value and do not correspond to our analysis (Poltzer-Ahles & Piccinini, 2018). To explore this relationship descriptively, we can perform a Median Split on the other contribution variable. For the high other-contribution set, abstract concepts had an average IOS of  $M = 3.46$  compared to  $M = 3.40$  for concrete. This pattern was reversed for the low other-contribution set, for which concrete concepts had a higher average IOS of  $M = 3.38$  than abstract concepts  $M = 3.22$ .



**Figure 6.** To visualize the descriptive pattern, the other contribution variable was split into four equal-sized bins; the error bars represent 95% confidence intervals of the mean within each category.

The ordinal mixed model analysis revealed an interaction between category (abstract versus concrete) and other-contribution with respect to the IOS dependent variable, with a positive coefficient (+0.06, SE = 0.03) indicating that for higher other contribution values, specifically abstract concepts, led participant to respond with higher iOS values. This coefficient was close to, but did not include zero, with a 95% credible interval:[0.01, 0.11]. The posterior probability of the coefficient being positive was high,  $p = .99$ . However, given the small effect size suggested by the descriptive averages and the pattern seen in Figure 6, we decided to also compare the model with the interaction against the null model without the interaction using Bayes factors, which yielded  $BF = 0.61$ , which is at most anecdotal evidence for the alternative model against the null model.  $R^2$  values of the model with the interaction were 1.2% higher than the model without the interaction, and the 95% credible intervals of these  $R^2$  values overlapped, which suggests that it is plausible that the interaction does not explain the variance. This model also includes difficulty as a covariate. There was no main effect of concept type (abstract versus concrete) on iOS, with only a numerical trend for abstract concepts leading to lower iOS scores (logit coefficient = -0.04, SE = 0.01, 95% CrI [-0.06, -0.02]). The posterior probability of this effect being of the same sign was 0.91. There was also no indication of a strong main effect of abstract concepts on iOS (-0.21, SE = 0.52, CrI [-1.26, 0.81]), with the posterior probability of this effect being positive,  $p = .67$ . The other main dependent variable, physical-interpersonal closeness, also did not yield strong evidence for an interaction between other contribution and concept category (logit estimate = 0.01, SE < 0.005, Cr [0.00, 0.02]), although a clear trend. The posterior probability of the interaction coefficient being

positive was  $p = .94$ . There was similarly no sign of any main effects, with the coefficient firmly included zero ( $-0.05$ ,  $SE = 0.10$ , 95% CrI:  $[-0.25, +0.15]$ )

### **Interim Discussion Experiment 1**

Experiment 1 shows that participants feel psychologically and physically closer during the interactive than the non-interactive condition. The results align with previous literature (Aron et al., 1997; Candini et al., 2017; Gessaroli, Santelli, di Pellegrino, & Frassinetti, 2013; Schmitz et al., 2023), and confirm the value of collaborative exchanges, like an online written synchronous CMC, in modulating the perceived interpersonal space and the perceived overlapping between self-other representations (Aron, Aron & Smollan, 1992; Shafaei et al., 2020). When interacting with a new person through a social platform, individuals engage in an inferential process to predict the other's identity, personality, opinions, values, based on conversational content and emotional cues. This interpersonal inferential mechanism seems to be sufficient to allow the establishment of a connection in humans (Aron et al., 1997), expressed by a sense of closeness both at a psychological and embodied-physical level. In addition, results show that linguistic interactions promote a sense of closeness already after five minutes of written synchronous CMC. Interestingly, the psychological closeness increased proportionally with the perception of how pleasant and intimate the exchange was perceived and decreased proportionally with how difficult the same was perceived. The perceived difficulty refers to the challenge of co-building a productive and coordinated new dialogical experience grounded on similar/different perspectives about the conversational topic. The more a dialogue appears difficult, the more people implicitly experience a sort of impasse or communicative failure regarding conceptual negotiation, self-disclosure, and autobiographical expressions, all of which impact the richness of the interaction. Interpersonal distance (Lisi et al., 2021; Fini et al., 2021), which is a more embodied and physical measure and concerns the imagined comfort zone of another person, has been found to be modulated by pleasantness and perceived difficulty, following a similar pattern as the IOS scale. Previous research demonstrated that the pleasantness of an interaction predicts the interpersonal distance between two interactors (Lahnakoski et al., 2020) and that interpersonal distance is modified by the quality of the social interaction and by positive/collaborative experiences (Candini et al., 2017; Gessaroli, Santelli, di Pellegrino, & Frassinetti, 2013). Finally, regarding the abstract/concrete topic of the conversation, Experiment 1 did not yield conclusive evidence for the hypothesis that abstract concepts, requiring more the other's contribution, lead to an increased closeness. However, we only

sampled eight concrete and eight abstract concepts, which, given the diversity of abstract concepts (Muraki et al., 2020; Villani et al., 2019), could mean that we have dealt with a specific sample. We therefore decided to sample a larger variety of concepts in Experiment 2, as well as improve the precision of our inferences by also sampling a larger group of participants. Experiment 2 also improves on Experiment 1 by performing a subgroup analysis (Villani et al., 2019) for specific types of abstract concepts

## **Experiment 2**

Experiment 1 provides anecdotal evidence about the hypothesis that the more actors perceive interlocutors' contribution to a conversation about abstract topics, the more psychologically close to the interlocutors they feel. The current study expands the previous one, by exploring whether conversing about different kinds of abstract concepts (Philosophical/Spiritual (PS), Inner/Emotional (EM), Self/Sociality (SS), and Physical/Spatio-Temporal/Quantitative (PSTQ), see Villani et al., 2019) might differently modulate psychological and physical/interpersonal closeness. The experimental paradigm is identical to the one of the study described above, except for the fact that the present one included only the interactive experimental condition. In this regard, a recent study has shown that during simulations of conversational exchanges, in which participants responded to sentences involving different sub-kinds of concrete (i.e., Tools, Animals, Food) and abstract concepts (i.e., Philosophical/Spiritual, Inner/Emotional, Self/Sociality, and Physical/Spatio-Temporal/Quantitative), the content of the verbal production changed (Villani et al., 2022). Specifically, abstract concepts generated higher uncertainty and more interactive exchanges than concrete ones. Along with these results, our aim here was to better explore how prompting conversations with concepts of different nature might affect the sense of closeness between the interlocutors. As declared in the preregistration, we expect to find that the higher a participant rates the Other Contribution to the conversation, the higher is the interpersonal closeness between the interlocutors, particularly with abstract concepts, but not with concrete ones. Moreover, we expect to find that the Other Contribution will impact the psychological/physical interpersonal closeness depending on the sub kind of abstract concepts. Experiment 2 sampled a wider range of concepts from different subgroups of abstract concepts. Moreover, given indications of a weak effect in Experiment 1, we increased our sample size based on a power analysis.

### **Pre-registration**

Experiment 2 hypotheses, methods, and analyses were formally pre-registered (<https://osf.io/ynuew>). As declared in the preregistration, here we aim to provide stronger evidence about the trend found in Experiment 1—the more actors perceive interlocutors' contribution to a conversation about abstract topics, the more psychologically close to the interlocutors they feel—and to explore the modulation of IOS and interpersonal-physical distance by different sub clusters of abstract concepts. We have to clarify that, although in the preregistration we declared that for data analysis it would have been adopted a frequentist approach by applying Linear Mixed Model (LMM) with the maximal random structure, considering that many models were not converging, we decided to shift to a Bayesian approach.

## **Participants**

One-hundred and thirty-six participants were recruited for the study. Six participants were excluded from the Interactive task as outliers (see 'data analysis' section), so that the final sample included one hundred and thirty participants (60 males, 77 females, group average age = 25 years; group average years of education = 16). The sample size of 130 participants is based on a power analysis conducted with the software MorePower 6.0.4 (Campbell & Thompson, 2012) for a five level, one factor within participants design. The analysis indicates that the sample size is adequate for detecting a small to moderate effect size  $\eta^2 = .035$ , with a power of .95 and an alpha level (two-sides) of .05. The experimental protocol was approved by the ethics committee of the Department of Dynamic, Clinical Psychology and Health Studies, Sapienza University of Rome (Prot. n. 0001040 16/11/2020), and was carried out in accordance with the ethical standards of the 1964 Declaration of Helsinki and later amendments. Participants gave their written informed consent to take part in the study and were naïve as to the purpose of the experiment.

## **Stimuli**

The stimuli consisted of sixteen abstract (Beginning, Thrill, Competition, Sum, Enigma, Fame, Cause, Trend, Discovery, Oath, Habit, Attention, Trick, Tale, Silence, Pride) and sixteen concrete concepts (Fisherman, Mirror, Bookcase, Pumpkin, Tailor, Uniform, Oak, Tourist, Aeroplane, Cock, Table, Bottle, Notebook, Pot, Cactus, Salad). Abstract concepts were selected from the Italian database by Villani and colleagues (2019); concrete concepts were instead extracted from Della Rosa et al.'s (2010) and Montefinese et al.'s (2014) Italian databases. These stimuli were selected through a previous validation survey in which 112 participants, divided in four groups, were asked to evaluate on a 7-point Likert scale, 149 abstract and 157 concrete concepts (divided in four

different surveys) in terms of valence. Concrete concepts were divided a posteriori in two macro categories: “organic” and “inorganic” and we performed a Linear Mixed Model featuring as fixed factor Subclusters (concrete organic, concrete inorganic, Philosophical-Spiritual, Inner/Emotional, Self/Sociality, and Physical/Spatio-Temporal/quantitative), as dependent variable Valence ratings and as random intercepts Participants and Words. The model did not yield a significant effect of Subcluster,  $F(4, 24.53) = 0.6068$ ,  $p = .662$ ), showing that subclusters did not differ among each other.

### **Online Chat Task**

The experimental procedure was the same as Experiment 1, with the only difference that here there was only the Interactive condition

### **Statistical analysis**

The statistical analysis for Experiment 2 largely mirrored that of Experiment 1. One noteworthy change is that we additionally performed a subgroup analysis on the abstract concepts only, looking at differences between the four subclusters of abstract concepts. All models include perceived difficulty as a covariate

### **Results**

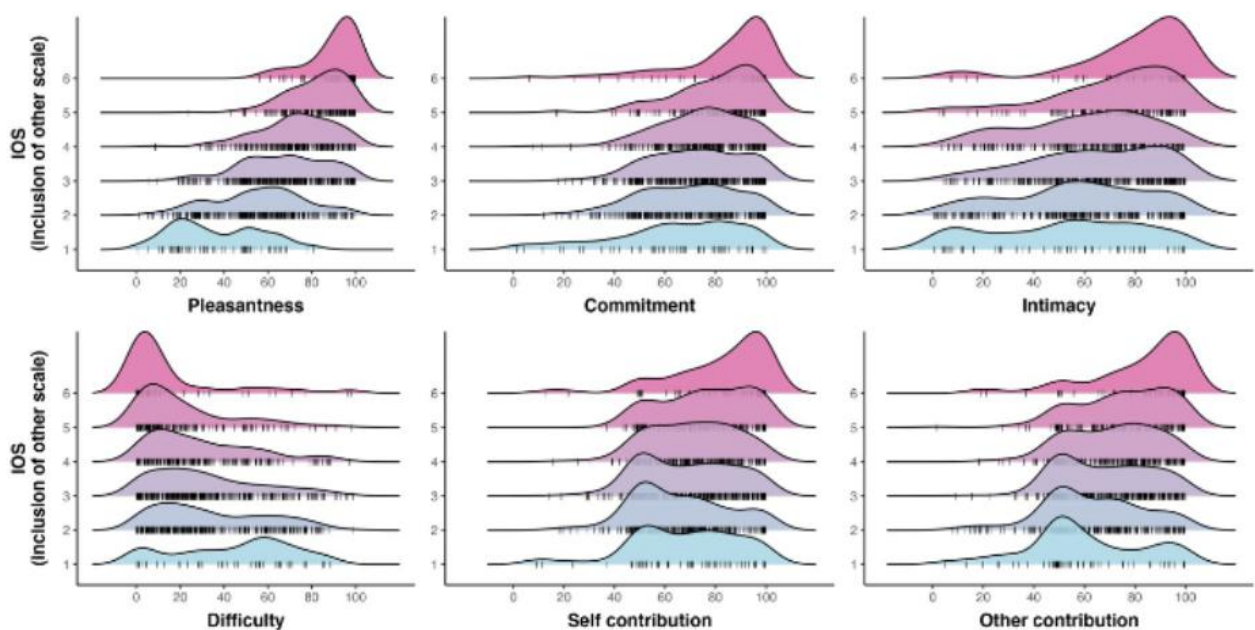
#### *Covariate analysis*

Table 2 shows Pearson's correlations between all covariates. The highest correlation was between ‘self-contribution’ and ‘other-contribution,’ as well as between ‘commitment’ and ‘self-contribution’ (both:  $r = 0.64$ ). Following this, the next highest correlations were obtained between ‘pleasantness’ and ‘other-contribution’ ( $r = 0.59$ ) and ‘pleasantness’ and ‘self-contribution’ ( $r = 0.45$ ). As was the case with Experiment 1, the ‘difficulty’ variable was negatively correlated with ‘pleasantness’ ( $r = -0.37$ ), as well as more weakly so with all the other variables.

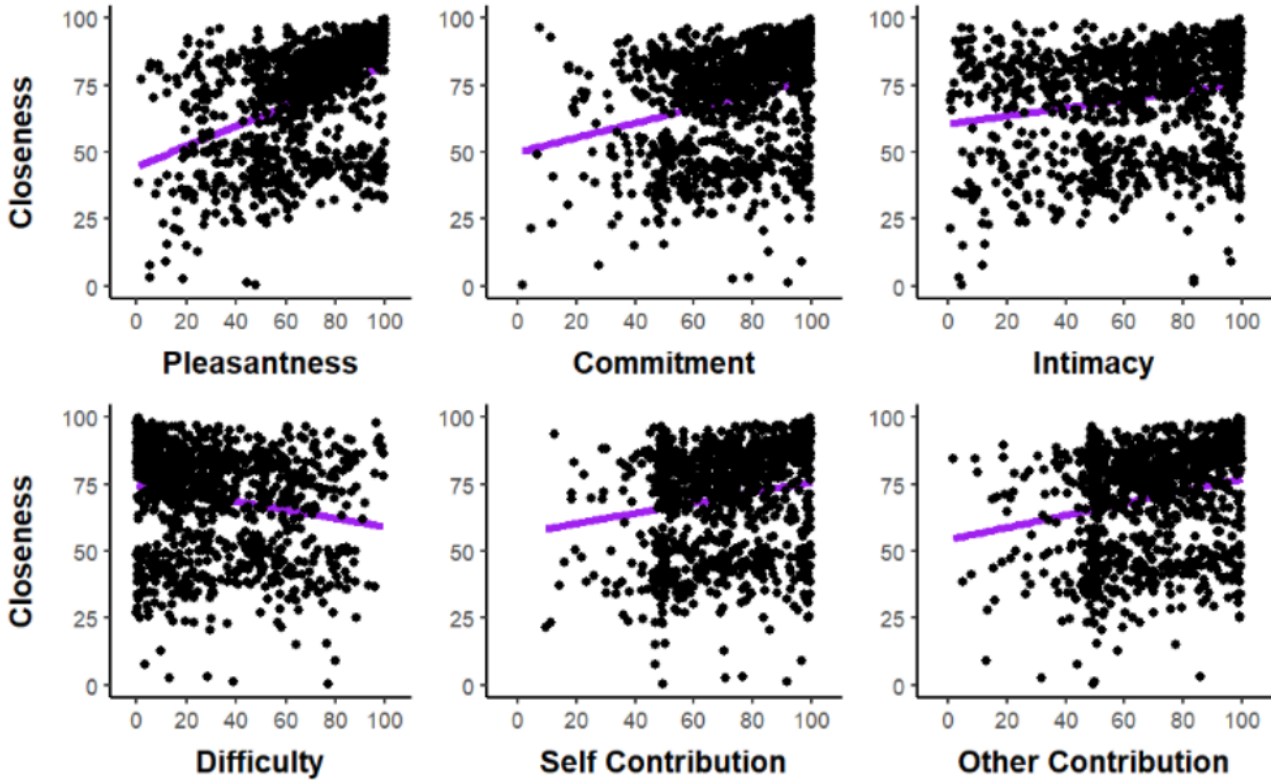
**Table 2.** Pearson’s correlations between all the covariates, i.e., Pleasantness, Commitment, Intimacy, Difficulty, Self-Contribution and Other-Contribution.

	pleasantness	commitment	intimacy	difficulty	self-contribution	other-contribution
pleasantness		0.54	0.53	-0.37	0.45	0.59
commitment			0.53	-0.18	0.64	0.54
intimacy				-0.23	0.44	0.43
difficulty					-0.21	-0.19
self-contribution						0.64
other-contribution						

The overall pattern of how these covariates behaved with respect to iOS was similar to Experiment 1, as shown in Figure 7. A mixed logistic ordinal model with iOS regressed onto all covariates showed a positive effect of pleasantness (logit coefficient = 0.09, SE = 0.01), with a 95% credible interval far away from zero, [0.08, 0.11], that was associated with a high posterior probability of being positive  $p = 1.0$ . On top of this, intimacy was positively associated with iOS (+0.01, SE < 0.05, 95% CrI: [0.00, 0.92],  $p > 0 = .98$ ), and difficulty was negatively associated with iOS (-0.01, SE < 0.005, 95% CrI: [-0.02, 0.00],  $p < 0 = .99$ ). Similar results were obtained for interpersonal closeness in Figure 8.

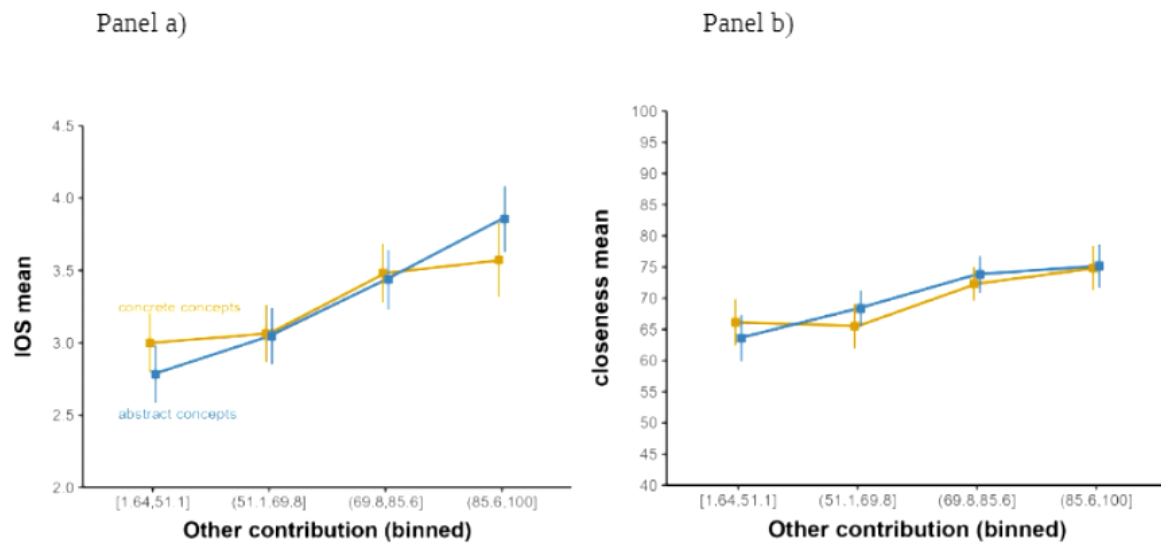


**Figure 7.** ‘Ridge plots’ showing how all six covariates differ for different levels of the Inclusion of Other Scale (IOS).



**Figure 8.** ‘Regression plots’ showing how all six covariates differ for different levels of interpersonal closeness (VAS).

The main hypothesis, an interaction between category and Other-contribution on iOS, was not confirmed. Figure 9 provides a visualization of the absence of an interaction between other contribution (x-axis) and concept category (separate lines).



**Figure 9.** To visualize the descriptive pattern, the other contribution variable was split into four equal-sized bins; the error bars represent 95% confidence intervals of the mean within each category.

For the mixed logistic ordinal model, the coefficient of the interaction was 0.0, with a 95% credible interval centered on zero: [-0.02, +0.03]. The Bayes factor comparing the model with the interaction against the model without the interaction indicated support for the null hypothesis  $BF = 0.07$ . The two models were also equivalent in  $R^2$ . Figure 9 shows that in the Experiment 2, there was a clear trend with higher other contribution values leading to higher iOS values. This was reflected in the model in a positive coefficient of other contribution (+0.06,  $SE = 0.01$ ), with a 95% credible interval far away from zero: [0.04, 0.08], posterior probability of being positive  $p > 0 = 1.0$ . The difficulty covariate was also negatively associated with iOS (-0.03,  $SE < 0.005$ ), with a 95% credible interval that did not overlap with zero, [-0.04, -0.02], and a high posterior probability of being negative  $p < 0 = 1.0$ . As was the case in Experiment 1, there was no indication of a main effect of concept category (+0.02,  $SE = 0.22$ ), with a 95% credible interval firmly entered at zero: [-0.40, +0.45],  $p > 0 = 0.55$ ). Similar results were obtained for the model with interpersonal closeness as a main dependent measure, as well as with interactions with concept category and self-contribution on either iOS or interpersonal closeness (all posterior probabilities of the interaction coefficient being of the same sign 0.20-0.80). Zooming in to the subgroups of abstract concepts, the pattern was similar, in that there was no indication of any interaction effects for other contribution for any of the four subgroups of abstract concepts. All coefficients were close to zero (between -0.01 and +0.01) and associated with large standard errors ( $SE = 0.02-0.03$ ) and 95% credible intervals firmly centered on zero. Posterior probabilities of coefficients of the interaction coefficients being of the

same sign ranged from 0.56 to 0.7, indicating that there was no compelling indication for any coefficient reliably excluding zero. The same was the case for the mixed beta regression model with interpersonal closeness as a dependent variable, where posterior probabilities of the interaction coefficients being of the same sign ranged from 0.64 to 0.71, all within the range of probabilities that do not firmly exclude zero. The same was the case when interactions were entered with self-contribution rather than other contribution, for iOS as a dependent measure (posterior probabilities of interaction coefficients being of the same sign = 0.55-0.79) and interpersonal closeness as dependent measure (posterior probabilities = 0.67-0.85).

### **Interim Discussion Experiment 2**

Contrary to our expectations, results of Experiment 2 suggest that psychological closeness (IOS) does not depend on the degree of abstractness of the conversational topic regardless of which subgroup of abstract concepts was considered. The same theoretical conclusions are furthermore obtained if we use a measure of interpersonal closeness. The Words as Social Tool (WAT) theory hypothesizes that abstract concepts, compared with concrete ones, are acquired through socio-linguistic interactions, which might represent the scaffolding for abstract concepts acquisition. We then hypothesized that the necessity to reduce the uncertainty and to negotiate the meaning (see Mazzuca et al., 2022; Mazzuca & Santarelli, 2022; Villani et al., 2022) might impact interpersonal dynamics. Our results, however, do not provide compelling evidence about the involvement of abstractness degree in interpersonal dynamics (e.g., IOS and interpersonal closeness).

The reason for this lack of significant results may be different. The first one is that we did not actually look into the conversations- this work will be the object of follow-up study - thus we do not know whether the body of the conversation prompted by an abstract as compared with a concrete one, holds also a more abstract level. Second, we did not apply any manipulations to control the abstractness level of conversation, for example stressing a common and “cultural” goal to achieve, which might ask for a deep understanding of conversational topics and thus prompt less Interactive/spontaneous and more learning-centered interactions. Third, to allow the realization of a similar subtle interactive process the written synchronous CMC should have lasted more than five minutes; this may have helped to potentiate eventual discussions, negotiation and opinions exchanges and would have been consistently expressed by the sense of closeness as a function of the conversational topic. Nevertheless, in Experiment 2 it emerges that, regardless of the abstractness degree, the more participants considered as important the other’s contribution to the conversation, the more they perceived themselves as psychologically and physically close to the

other interlocutor. In other words, recognizing that the other plays a crucial role in linguistic exchanges leads to perceiving them as closer both psychologically and physical level. Finally, the same results identified in Experiment 1 are replicated also in Experiment 2: the psychological closeness increased proportionally with the perception of how pleasant and intimate the exchange was perceived, and decreased proportionally with how difficult it was perceived.

### **3.2.2. General Discussion**

#### *Digital connection and sense of closeness*

In two experiments, we showed that online interactive written chats, lasting less than five minutes, are sufficient to increase the interpersonal-psychological and physical closeness between two interlocutors. We asked dyads to chat about a specific topic with another individual (interactive condition), or to write about the same topic through a social network (ICQ), knowing that on the other side of the screen, another person was doing the same and that she/he would have read the text later (non-interactive condition). In the interactive condition, participants felt both psychologically and physically closer than in the non-interactive condition. Our results shed light on the powerful effects of brief on-line written conversations in modulating the sense of closeness between two interlocutors. In the last twenty years, communicating has become more than ever a virtual practice. Such an innovative modality to stay in relation with others largely impacted our socio-emotional sphere and started a deep process of transformation of our mental interpersonal representations. A bunch of research has already been conducted on how CMC technology allows easy interaction and has the power to recalibrate the psychological distance— as it is intended by the Construal-level-theory (Trope, Liberman & Stephan 2007) —among remote interactors (Oh, Curley, & Subramani, 2008; Norman, Tjomsland, & Huegel, 2016; see also Wilson 2005). Thinking about a person as spatially distant leads people to think of them as socially distant (Liberman, Trope, & Stephan, 2007). People who talk over large spatial distances treat each other with greater formality, less familiarity, and act less cooperatively than people talking over short distances (Bradner & Mark, 2002; Stephan, Liberman, & Trope, 2010, 2011). It is known that mediated communication shapes perceptions of spatial distance (Andrade, 2014; Won et al., 2018) and promotes the so-called “death of distance” phenomena (Cairncross 2001): by bridging physical distance, mediated communication can decrease social distance (Oh, Curley, & Subramani, 2008, see also Wilson 2005). Not only do technological devices tout court allow cancelling the distance among the interactors, but also the interaction style itself might contribute to enhancing the sense of closeness:

in order to compensate for the absence of non-verbal language, CMC interactors are more direct and allow more intimate uncertainty reduction behaviours than face to face interactors (Tidwell & Walther, 2002). Within the umbrella term of Computer Mediated Communication (CMC), we focused on exchanges composed of a few turn-takings, in other words on instant messaging (IM) and on its short-term effects on interpersonal distance. Specifically, here we examined the perception of interpersonal distances among the interactors. The psychological interpersonal distance refers to the degree of perceived overlap between the self and other mental representation (Inclusion of other in self-scale, iOS, Aron & Aron, 1992) and the physical interpersonal distance (measured through a Visual Analogue Scale–VAS– scale) refers to the representation of an hypothetical comfort spatial space between the self and the other. While iOS can be conceptualized as less embodied and a more emotional measure, the physical-interpersonal distance is an embodied measure, linked to the representation of the shared social space (Hall, 1966; Bzdok & Dunbar, 2020). Both measures indicate that when two individuals chat in an “aspatial context”, like an on-line written chat, they represent themselves as closer compared to when they know about their reciprocal presence and write synchronously about the same topic, knowing that the other will read the text later. Thus, online interactivity (see Norman, Tjomsland, & Huegel, 2016) and participating in a dialogue appear to determine an increased sense of closeness. Importantly, conversing about a topic is a form of joint action (Clark, 1996; Garrod & Pickering, 2004) and requires a collaborative attitude to lead the exchange towards accomplishing the shared goal: building together a new and rich discourse. If we conceptualize the conversation, including the CMC written synchronous exchanges, as a collaborative task, our results extend to a virtual environment what has been already found during interpersonal sensorimotor interactions: imaging to collaborate on a common task increases interpersonal closeness between the interactors more than performing parallel actions (Schmitz & Reader, 2023). Importantly, studies highlighted that increases in physical distance between individuals and groups are inversely linked to factors such as interpersonal liking, enjoyment, and communication frequency (Wilson et al. 2005), similarly, the pleasantness of interaction predicted the interpersonal distance among the interactors in a joint action task (Lahnakoski et al., 2020). In keeping with this evidence, we found that the more the conversation was considered as pleasant, the more people felt psychologically and physically closer to the interlocutor. Moreover, the more the conversation was considered as intimate in experiment 1, the more interlocutors felt psychologically closer, while, in experiment 2 they felt both physically and psychologically closer. Intimacy (Argyle and Dean 1965) is related to physical proximity and the topic of conversation, and it is associated with self-disclosure. It is already known in the literature

that mediated communication can even lead to greater intimacy than face-to-face communication (Walther, 1996); for example, Bazarova (2012) demonstrated that private disclosure prompted greater inferences of relational intimacy during on-line chats. There is evidence that, as compared with face-to-face conversations, synchronous CMC might provide a unique context to compensate for the absence of nonverbal cues, as well as identity cues. According to the hyper personal perspective of CMC, these factors can contribute to the enactment of more intimate exchanges compared to their face-to-face counterparts (see the hyper personal perspective of CMC, Tidwell & Walther, 2002). We also found that the more a conversation was perceived as difficult, the more the interpersonal distance between the interlocutors increased. Evaluating a dialogue as difficult, entails admitting that any collaborative efforts to produce a common knowledge did not succeed and probably points to the failure of the linguistic exchange, explaining, thus, the increases in interpersonal distance between the interlocutors, both physically and physiologically. In experiment 2, also another relevant result has been highlighted: the more participants thought that the other interlocutors contributed to the conversation, the more they felt psychologically and physically close to them. Recognizing that the other person has actively participated in the dialogue, modulates interpersonal closeness and suggests the role of the spontaneous attribution of merit in reaching a shared goal: co-building knowledge can improve interpersonal closeness. Regarding the main pre-registered hypothesis of Experiment 2, that the abstractness of the topics might modulate both the psychological and physical closeness, we did not provide conclusive evidence. The Words as Social tool—WAT— theory proposes that language and social interaction are more crucial to learning abstract concepts and that, due to their indeterminate meaning, abstract concepts promote social exchange and negotiation more than concrete concepts (Borghi et al., 2018 a,b, 2022; Fini & Borghi, 2019; Fini & Era, 2021). Starting a conversation with an abstract concept might have produced higher levels of negotiation and inferential process about the other's perspective, and consequently, an increased collaborative attitude permeating the dialogue, leading to interpersonal proximity. Experiment 1 provided inconclusive evidence for the role of abstract concepts in leading to higher judgments of other-inclusion, as a function of how much participants recognized the other's contribution to the conversation, but a numerical trend in that direction. In Experiment 2 we did not replicate the trend identified in Experiment 1 and we did not find a specific involvement of different subclusters of abstract concepts sampled from Villani et al., (2019).

### *Implications for abstractness*

The results have many implications for theories on abstractness. In particular, they are very informative for theories according to which social interaction plays a major role for abstract concepts, such as the WAT theory. The WAT theory is a relatively new proposal. One of its central tenets (for details, Borghi et al., 2023) concerns social interaction's role during conceptual acquisition and processing. Because of the unique character of abstract concepts, particularly the fact that they collect heterogeneous exemplars, hardly imageable and not perceptually similar, it proposes that people need others more during the acquisition of abstract than concrete concepts. Others would offer linguistic explanations and social support to learn them. Some preliminary evidence favouring this shows that, during abstract concept learning, both in children and adults, the mouth motor system might be more recruited than the hand motor system (Borghi et al. 2011; Granito et al. 2015; Barca et al. 2017, 2020; Reggin et al. 2021), supporting the importance of language for abstract concepts acquisition. According to WAT, when children and adults learn new words, they verbally rehearse possible words and use inner dialogue to arrive at the correct meaning (inner social metacognition) or prepare ourselves to revert to others to ask for it (social metacognition). A similar mechanism also characterizes guessing abstract words starting from images. Likely because of the lower imageability of abstract concepts, participants would need others more to help them guess which abstract concepts images refer to. This leads participants to resonate more with the people who assist them in matching words and images (Fini & Era, 2021). Evidence supporting this also comes from ratings showing that people think they need others more to learn abstract concepts (Villani et al., 2019). WAT also proposed that people need others more during the processing and use of abstract concepts. Because of their fuzzy boundaries and indeterminate character, people might negotiate their underspecified word meanings with others, as when scientists debate to arrive at a new concept's definition. More simply, they might seek to understand better what others intend. Notably, here the interaction between the people involved is symmetrical rather than asymmetrical. Here evidence on the role of social interaction is more scattered. First, rating studies showed that people feel less confident knowing abstract concepts' meanings (Mazzuca et al., 2022). Second, in simulated conversations, participants express more uncertainties and ask more questions, and this is particularly true for the more abstract concepts, philosophical-spiritual ones (Villani et al., 2022). Third, ratings indicated that it is easier to start a conversation using abstract concepts, particularly with self-sociality concepts (Fini et al., 2023b). The present study shows that, in Interactive conversations, starting from an abstract rather than a concrete concept does not lead to differences in perceived closeness with others. Why? The reasons can be many. First, it is possible that, being conversation a form of joint action, it enhances

closeness in any case, independent of the involved topic (unless in cases of conflict or different opinions). Another is that during an Interactive conversation, which is not aimed at a specific goal (e.g., learning a new term, defining together a term, preparing actions based on the term knowledge, etc.), there is no need to reach a deep mutual understanding of the concept. Third, it is possible that, even if the prompt for the conversation is a concrete vs. abstract concept, the flow of the conversation directed elsewhere. Careful analyses of the conversation content will allow to discover whether this is the case. Finally, it is possible that because the interaction is symmetric, the other is not crucial to enrich knowledge differently in the case of a concrete or an abstract concept. Further reasons for the absent result are possible. For example, we might be able to capture the need of others with implicit but not explicit measures, such as the request to express a feeling of closeness. Overall, these results cast doubt on the claim of WAT, according to which we would need others not during the acquisition but in the everyday use of abstract concepts in on-line written brief symmetric relations. Further analyses and novel paradigms are necessary to test whether we need others more with abstract than concrete concepts in symmetric relationships.

### **Limitations of the study**

In the current study, some limitations can be identified regarding the verification of the hypothesis about the modulation of the interpersonal dynamics between the interlocutors as a function of the abstractness of the conversational topic. First, the quality of the interaction between the interlocutors was spontaneous and in order to preserve that character we did not control for the physiological content fluctuation typical of dialogical exchanges. By adopting such an ecological approach, we might have underestimated the need to exert manipulation on the conversational content. We did not control for the extension of the abstractness level of the initial topic to the entire body of the conversation. Second, the interactions, being so ecological, were less goal-directed and not performant; in other words, a deep understanding of the topic was not required in order to later re-explain the same to another person. For these reasons, the other's contribution to the conversation probably assumed a less utilitarian value: the goal to accomplish was not a "cultural/divulgate" one but the pure enjoyment of the conversation. Undoubtedly, the study suggests that conversations might represent diversified terrains of emotional and metacognitive experiences, allowing the speculation that some metacognitive experiences (i.e., other contribution, self-contribution, commitment) might strongly impact interpersonal dynamics in more performant-interactive scenarios as compared with spontaneous exchanges.

## **Conclusions**

When conversing online we are highly sensitive to the pleasantness, the difficulty, the intimacy, and the other's contribution to the dialogue. All these dimensions impact mental representations: people feel psychologically and physically closer when chatting compared to when they are just digitally connected with someone. Importantly, the subjective perception of the quality of the dialogue modulates the sense of closeness, while the topic of conversation is less relevant. Thus, we can conclude that CMC technology increases connection, but connection is not enough to feel authentically connected or closer to someone; what matters is the dialogue and the subjective perception of its quality. Knowing the importance of talking online and of conducting a qualitative dialogue, even if for a few minutes, can be particularly relevant to ameliorate the quality of life in contexts characterized by social isolation, mental illness, sickness, and for older people or hospitalized dwellers that cannot spend too much time in real social interactions due to their health conditions, but that can still benefit from this low-cost modality to feel closer and less alone.

### 3.4. Summary

Chapter 3 aimed to provide a comprehensive examination of the intricate relationship between language and social cognition in interactions. Each section aimed at contributing to a deeper understanding of how language, interpersonal representation processing and emotional engagement influence the way individuals interact with one another.

The exploration of conversation initiation through abstract concepts (ACs) and concrete concepts (CCs) in Chapter 3.1. provides a foundation for understanding how conceptual frameworks can influence social interactions. The findings indicated that ACs are generally perceived as easier to discuss, especially among individuals unfamiliar with a specific topic. This aligns with previous research suggesting that ACs requires more cognitive processing and social engagement (Borghetti et al., 2019a,b). Specifically, the emotional-experiential and semantic-metacognitive dimensions identified as influencing conversational ease support the notion that individuals perceive conversations of ACs as stimulating interactions and negotiations (Mazzuca et al., 2022; Villani et al., 2022). Chapter 3.2. focused on the dynamics of digital communication on ACs and CCs and their differential effect on perceived closeness between participants. The results revealed that interactive chats, regardless of whether they involved ACs or CCs, significantly enhanced feelings of psychological closeness. This finding highlights the importance of interactive and participatory elements in digital communication. Notably, the absence of a significant difference between the types of concepts used in fostering closeness suggests that the quality of interaction and the emotional tone of the conversation may play a more critical role than the nature of the concepts themselves.

Together, these research illustrate a nuanced relationship between concept types and social interaction. While ACs may facilitate easier initiation of conversations by engaging participants cognitively and emotionally, the actual process of communication—particularly in digital settings—seems to hinge more on interaction dynamics rather than the abstract or concrete nature of the topic. This indicates that in digital environments, fostering interactivity and intimacy can be key to enhancing social bonds, irrespective of the concept types being discussed. These insights underscore the importance of understanding both conceptual frameworks and communication styles in promoting effective social interactions and the potential for ACs to stimulate the creation of deeper conversations in various contexts.

## 4

# General Discussions

The present thesis, as stated in the introduction, aimed at investigating the interplay between social interactions, sensory integration, communication and social cognition dynamics. It was structured into two main experimental chapters: Chapter 2, focusing on how interpersonal motor interactions affect multisensory and motor processing, explored how visual and tactile information are integrated during and after joint actions and whether interpersonal motor interactions affect individuals' body ownership, but also how social factors, such as racial biases, can affect motor coordination and interactions; Chapter 3 shifted the focus to linguistic communication, particularly on how conversating on abstract concepts (ACs) compared to concrete concepts (CCs) can contribute to the ease of starting conversations and on whether online interactions can enhance feelings of interpersonal closeness.

Chapter 2.1. explored how interpersonal motor interactions influence the integration of visual and tactile information. Consistent with previous findings on visuo-tactile incongruency effects (Brozzoli et al., 2009; 2010; Senna et al., 2019; Patanè et al., 2021), this research demonstrated that VTI is modulated when visual cues appear on an individual's target object but also when they appear on a partner's object, but only in interactions that require continuous prediction of the partner's movements. This suggests that interpersonal motor interactions engage neural processes linked to multisensory integration, implying that the partner's movements and objects become instrumental for the agent's action goals. This process mirrors action-based somatosensory resonance observed in earlier studies (di Pellegrino et al., 1992; Rizzolatti & Sinigaglia, 2016), and supports the notion that in social contexts PPS functions as a multisensory-motor interface guiding social interactions. Notably, VTI effects emerged only when participants had to predict their partner's movements, emphasizing the importance of predictive mechanisms in joint actions and space representation mechanisms. These findings deepen our understanding of multisensory integration during social tasks and open avenues for further research on the underlying neural mechanisms involved and on how social factors, such as individual characteristics, affect sensorimotor processes during interpersonal interactions. Overall, the study highlights how sociality

extends perception and action within PPS representations, shaping individual behavior in joint actions.

Chapter 2.2. investigated how interpersonal motor interactions modulate both behavioral (i.e., Reaction Times) and neural (i.e., EEG patterns) markers of cross-modal interpersonal integration, focusing on the integration of visual and tactile/proprioceptive information. Behavioral results showed that subjects experienced stronger interference when processing spatially incongruent stimuli after interpersonal interactions, particularly for tactile over visual information. Moreover, RTs significantly diminished when both visual and tactile channels were engaged during prior interactions (M-PV condition), suggesting optimal multisensory integration respectively in incongruent and congruent trials (Ernst, 2002; Noppeney, 2021). Neurophysiological data further supported these findings, with neural responses modulated as early as 74 ms post-stimulus, depending on prior interactions, particularly in the M-PV condition. This study supports cross-modal integration in an interpersonal context, highlighting how multisensory processing is shaped by joint actions and sensory coordination. These results underscore the dynamic nature of sensorimotor integration, with implications for understanding the neural underpinnings of social interactions.

The study of Chapter 2.3. reveals how multisensory visuo-tactile integration and body ownership are modulated in social contexts, using a novel 'social' rubber hand illusion (RHI) paradigm. Proprioceptive drift increased with task repetition, regardless of stroking synchronicity, suggesting that social proximity during the task may diminish the role of synchronous cues, contrary to classical RHI findings (Botvinick & Cohen, 1998). However, stroking synchronicity did significantly enhance the sense of ownership and location (SRHI-Q), consistent with prior literature (Longo et al., 2008; Tsakiris et al., 2010), and led to stronger embodiment effects even when the hand was in an implausible position. This finding highlights how mutual engagement can strengthen multisensory integration. Notably, VTI effects were not influenced by prior synchronous or asynchronous manipulations, suggesting distinct multisensory mechanisms for body ownership and the classical multisensory processing (Spence et al., 2000) or may reflect a shrinkage of the self space representation relative to another person (Fossataro et al., 2023). In general, these findings contribute to our understanding of how social interaction shapes body perception and call for further exploration of the behavioural and neural mechanisms underlying these processes.

In Chapter 2.4. the studies examined how racial biases influence motor coordination during joint actions, focusing on the role of the mPFC through cTBS. Participants exhibited slower RTs when interacting with the out-group avatar, indicating an impact of racial membership on motor

coordination, consistent and expanding prior findings on implicit racial biases (Sacheli et al., 2015). However, grasping asynchrony and movement time were primarily influenced by movement type rather than avatar ethnicity, suggesting that racial biases mainly affect reaction times in joint actions. Despite previous evidence linking mPFC to prejudice regulation (Amodio, 2014; Sellaro et al., 2015), the study found inconclusive evidence for its role in modulating racial bias during joint actions, possibly due to the absence of baseline bias in the sample (see AMP results). These findings highlight the complexity of social factors in influencing motor interactions and suggest that while social categorization influences RTs, motor coordination may involve more complex variables than social biases alone that need to be explored more in depth, also at the neural level.

Chapter 3.1 investigated the dynamics of conversation initiation with abstract concepts (ACs) compared to concrete concepts (CCs), revealing that concepts rated as more pleasant, familiar, and positively valenced facilitate conversation initiation. Specifically, ACs are perceived as easier to discuss than CCs, particularly when familiarity is low, highlighting their advantage in the Emotional–Social Communication (ESC) dimension. The principal component analysis (PCA) identified two key dimensions: an emotional–experiential dimension (encompassing Familiarity, Valence, and Pleasantness) and a semantic–metacognitive dimension (comprising Word Confidence (WC), Context Availability (CA), Social Meaning (SM), and Open Negotiation (ON)). Findings indicate that while both dimensions contribute to the ESC, emotional–experiential factors predominantly influence conversation initiation. The results suggest that ACs evoke greater uncertainty, necessitating more social interaction and collaborative meaning-making, thus enhancing their conversational potential. Notably, Self-Sociality concepts, which relate to social experiences, emerged as particularly effective in facilitating dialogue. This research aligns with the theoretical framework of social grounding (Borghetti et al., 2017, 2018, 2019), indicating that ACs possess inherent qualities that make them conducive to social interaction, transcending mere emotional valence. Future studies should explore these dynamics in real conversational contexts and consider environmental factors influencing conversational choices. Understanding these relationships may improve the selection of linguistic stimuli in studies on conversational dynamics and social interaction.

Chapter 3.2. explored how brief online written interactions affect interpersonal closeness. Across two experiments, results showed that interactive chats lasting less than five minutes significantly increased both psychological and physical closeness compared to non-interactive exchanges. These findings align with Construal-Level Theory (Trope et al., 2007), which posits that mediated communication can reduce perceived social and psychological distance between individuals.

Dialogue quality, including pleasantness, intimacy, and the perceived contribution of the interlocutor, was found to modulate closeness. Interestingly, the topic's abstractness (versus concreteness) had no significant effect on interpersonal closeness, challenging some aspects of the Words as Social Tools (WAT) theory (Borghi et al., 2019a,b). The study concludes that while digital connection can foster a sense of closeness, meaningful dialogue plays a more crucial role in truly connecting people. These findings highlight the importance of high-quality online conversations for improving social connectedness, this can be especially important for individuals facing isolation, illness, or limited real-life social interactions, and set the base to explore these dynamics in the long run, comparing different populations and exploring more deeply the factors influencing interpersonal online dynamics.

In conclusion, the research presented in Chapters 2 and 3 illustrates the profound impact of social interactions on sensorimotor and interpersonal dynamics. The connections established between these studies underpin the importance of considering both sensorimotor and cognitive dimensions in understanding social interactive behaviour. Overall, they contribute to a deeper comprehension of the multifaceted nature of social interactions and their implications for individual perception and action. Future research should continue to investigate these intersections, exploring how various interpersonal factors influence both sensory and motor integration in different social environments.

# Appendix

## A1. Overview of the Publication Status of Chapters in this Thesis

Chapter number and title	Original text (not published before)	Submitted: no feedback received	Submitted: revision requested or revision submitted	Accepted/published
1. General Introduction	X			
2.1. On-line Interpersonal Motor Interactions Modulate Visuo-Tactile Integration Mechanisms	X in preparation			
2.2. Dyadic interactions based on multiple sensorimotor modalities modulate the behavioral and electrocortical markers of interpersonal cross-modal processing	X in preparation			
2.3. Toward a “Social” Rubber Hand	X in preparation			
2.4. The Hidden Cost of Outgroup Coordination: Action Delays in Joint Actions.	X in preparation			
3.1. Breaking the ice in a conversation: abstract words prompt dialogs more easily than concrete ones.				X <i>Language and Cognition</i>
3.2. Digital Connection, Real Bonding: Brief OnlineChats Boost Interpersonal Closeness Regardless of the Conversational Topic				X <i>Helyon</i>
4. General Discussion	X			
A2. Preregistration “Processing Abstract and Concrete Concepts: the Case of Aphantasia”				X <i>OSF</i>

## **A2. Preregistration “ Processing Abstract and Concrete Concepts: the Case of Aphantasia”**

### **A2.1. Introduction**

When individuals are prompted to close their eyes and think about an object, they “use the mind’s eye” to visualize it internally. Seeing the object through “the mind’s eye” is a process known as mental imagery. Mental imagery is closely intertwined with many cognitive functions, such as visual memory (Albers et al., 2013; Keogh & Pearson, 2011, 2014), spatial navigation (Ghaem et al., 1997), language comprehension (Bergen, 2007; Zwaan et al., 2002) and it also plays a pivotal role in moral and altruistic decisions (Amit & Greene, 2012; Gaesser & Schacter, 2014). Imagery has been shown to activate the visual cortices. The imagined content may be decoded from the cortical activity using an encoding model based on low-level depictive visual features, such as spatial orientation and contrast (Naselaris et al., 2015). Mental imagery also holds clinical relevance. It is commonly employed in cognitive behavioral psychotherapy techniques such as imaginal exposure and rescripting (Arntz et al., 2007; Holmes et al., 2007; Pearson et al., 2015). Importantly, not all people are endowed with the same level of mental imagery abilities. “Aphantasia” refers to a condition of reduced or absent imagery (Zeman, 2010). This condition can be acquired after brain damage or be present from birth (congenital aphantasia - Zeman, 2015). One interesting question is to which degree mental imagery contributes to processing abstract and concrete concepts.

Following the hybrid and the multiple representation theories (see for recent review by Muraki et al., 2023), conceptual meanings are derived from linguistic and sensorimotor processes. In keeping with the Words As Social Tools proposal (WAT - Borghi et al., 2017; Borghi, 2023), concrete concepts evoke sensorimotor experiences more intensely and frequently than abstract concepts, for which linguistic and social experiences might be more crucial. Indeed, the referents of concrete concepts are more imageable than those of abstract concepts, which would activate the five senses to a lower extent and be more difficult to associate with single perceptually bounded physical referents (Borghi & Binkofsky, 2014). The so-called “Concreteness Effect” refers to the advantage in the processing and recall of concrete (i.e., “apple”) over abstract concepts (i.e., “justice”-Paivio, 1990; Schwanenflugel, 1991; Schwanenflugel et al., 1992). In lexical decision tasks, reaction times (RTs) are faster, and accuracy is higher for words with an easily imageable referent (Hargreaves,

2012). This evidence aligns with the classical Dual Coding Theory (e.g., Paivio, 1986; Paivio et al., 1968), according to which all concepts are represented through the verbal system, but only concrete concepts are directly encoded through images. For instance, the abstract concept of “religion” evokes images only through the mediation of concrete concepts such as “church” (Borghetti et al., 2017).

Following these assumptions:

1) Concrete and abstract concepts activate both sensorimotor and linguistic experiences, but to a different extent: sensorimotor (including visual) information has a significant weight in concrete concept representation, whereas linguistic information in abstract concept ones;

2) Concepts' imageability contributes to the Concreteness Effect.

In line with this, we might expect that the Concreteness Effect in Aphantasic individuals might be reduced when linguistic strategies of meaning retrieval are inhibited, i.e., when only visual cues are available in the conceptual processing.

Since Aphantasic individuals show reduced visual imagery abilities, they might not benefit from the advantage of accessing concrete over abstract concepts through the visual channel. Consequently, they might rely more on propositional, verbal strategies. In the current study, we aim to test whether individuals with aphantasia do not present the classical Concreteness Effect when the linguistic strategies are inhibited while processing abstract and concrete words.

## **A2.2. Materials & Methods**

### **Design plan**

2x2x2 mixed design for each experiment (visual search task and semantic relatedness judgment task), with reaction times and accuracy as dependent variables, Category of Word (abstract, concrete), and Type of Interference (verbal interference condition, visual interference condition) as within-participants factors, and Group (Aphantasic group, Non-aphantasic group) as a between-participants factor.

### **Stimuli selection**

We preliminarily selected 40 abstract and 40 concrete concepts matched for Familiarity ( $X^2(1) = 3.5441, p = .06$ ), i.e., the level of perceived personal experience with a concept (FAM - Barca et al., 2002; Falcinelli et al., 2023). The word set was extracted from a previous study investigating the semantic properties of concepts (Fini et al., 2021). Then, in two preliminary surveys implemented on the online platform of Google Form, 74 non-aphantasic participants (38 female, Mage = 37.00) were asked to verbally describe each concept of the word set by detailing the first visual image that came to their mind and that they considered representative of it. We ensured participants were not aphantasic through the administration of two questionnaires within the same survey, i.e., a revised version of the Vividness Of Visual Imagery Questionnaire (VVIQ, Marks, 1973; Zeman et al., 2015, 2020) and the Plymouth Sensory Imagery Questionnaire (PSI-Q, Andrade et al., 2014). In line with this, overall, one subject was excluded from the sample since the scores indicated the presence of aphantasic characteristics, according to the VVIQ and PSI-Q (Marks, 1973; Zeman et al., 2015, 2020; Andrade et al., 2014).

Once we collected visual images' verbal descriptions, we processed texts through a text mining analysis using 'panda' Python's package (McKinney et al., 2010) to identify the most representative descriptions for each concept. After cleaning the data (e.g., by removing symbols and stop words), we applied a Latent Semantic Analysis (LSA) technique in topic modeling to this dataset. We identified a topic for each concept through the 'TF-IDF features' (Pedregosa et al., 2011), and the most frequent three-word chunks in each topic were extracted. Then, based on the three-word chunks, specific images representing the mental image were implemented for each concept, employing the online application of CANVA (<https://www.canva.com>). A target image was created from each verbal chunk, which well-represented the mental image evoked by the concept.

Along with the targets, four distractor images were created per each. Distractors were realized by keeping the same primary elements of the target images to maintain the same graphic complexity (e.g., background, number of elements represented) and varying the associated concept's representativeness. Figure 1 shows an example of the target image and distractors associated with one concrete (Panel A) and one abstract concept (Panel B) from our word set.

**Figure 1. Panel A)** An example of images constructed for the concrete concept of "Rooster", showing how the distractors (1,2,3,4) would differ from the target image (5) only in the represented animal and not in complexity or number of included elements, e.g., the farm in the background, the countryside scenario. **Panel B)** An example of images constructed for the abstract concept of "Suggestion", showing how the distractors (1,2,3,4) would differ from the target image (5) only in the represented scene and not in complexity or number of included elements, e.g., there are two people in all images and the background is the same.

**A**

1



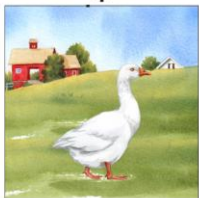
3



2



5



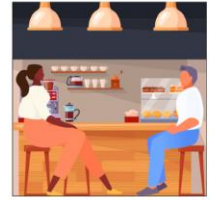
4

**B**

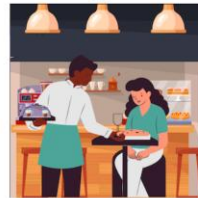
1



3



2



4



5

Then, to validate the selected images, the semantic association between one target image and four distractors to the main concept (40 abstract and 40 concrete concepts) was evaluated through two independent surveys administered to an independent sample of 42 non-aphantasic participants. More specifically, we asked participants: “How much do you think this image conveys the meaning of the concept?” and to provide their answers using a 7-point Likert scale (1 = “not at all”; 7 = “extremely”). We checked participants for aphantasic traits through the VVIQ and PSI-Q; three subjects were excluded from analyses since their questionnaire scores indicated the presence of aphantasic characteristics (Marks, 1973; Zeman et al., 2015, 2020; Andrade et al., 2014). We selected visual stimuli for the visual search and semantic relatedness judgment tasks based on the results. Specifically, we identified 48 concepts ( $n = 24$  abstract and  $n = 24$  concrete) and related images. They were randomly delivered through a between-tasks procedure (i.e.,  $N = 24$  words for each task, 12 of which were abstract and 12 concrete) to avoid the repetition of the same stimuli in both tasks that might bias participants’ responses. Linguistic stimuli resulted in being matched for familiarity,  $X^2(1) = 0.731$ ,  $p = .391$ , valence,  $X^2(1) = 1.967$ ,  $p = .161$ , pleasantness,  $X^2(1) = 2.194$ ,  $p = .139$ , figurative representativeness of the corresponding target images, estimate = .3661, SE = .4794,  $z$  value = .764,  $p = .445$ , and figurative representativeness of the four distractors, estimate = .1808, SE = .3739,  $z$  value = .484,  $p = 0.629$ .

## Tasks

### *Preliminary Image Presentation Task.*

Participants will be instructed to memorize a series of words displayed before each trial. They will be presented with eight lists, each comprising ten words. Their explicit objective will be remembering as many words as possible from each list. Following each list, participants will view and be asked to focus on the presentation of images associated with the relative concepts for three seconds (timings from Zhou et al., 2024). The images and concepts displayed will be only the target stimuli from the selected sample for the visual search and semantic relatedness tasks (see above). Upon completion of each trial, participants must recall the words they can remember from the previously displayed list. The experimenter will record the recalled words for each trial, and the procedure will proceed accordingly. This task aims to present the participants with the target stimuli passively. After this preliminary task, participants will perform the visual search and semantic relatedness tasks a week later to avoid learning effects.

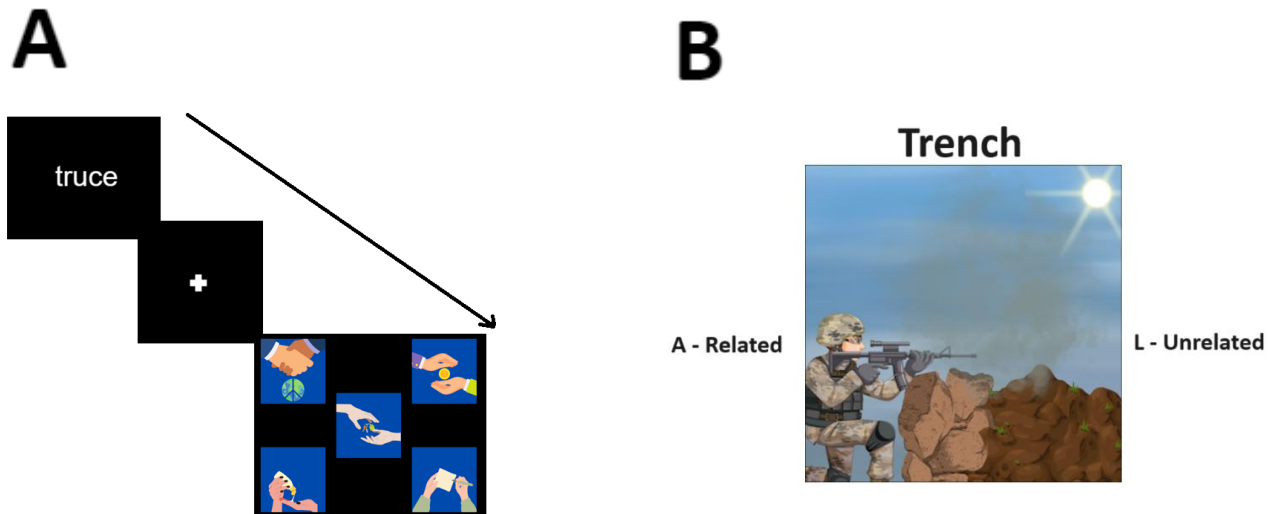
### *Visual search task.*

The visual search task will start by presenting the concept and the associated *ad hoc* images ( $n = 1$  target and  $n = 4$  distractors; see above). Then, within 4.1 seconds (for the timing choice, consult the pilot section below), participants will be asked to indicate the image that best represents the concept from the array. The stimuli position will be randomized across participants trial by trial. Participants will be asked to select the target image quickly and accurately (Figure 2, Panel A).

### *Semantic relatedness judgment task.*

The semantic relatedness judgment task will start by displaying the target or one of the distractors' images associated with the relative concept for 1.2 seconds (for the timing choice, consult the pilot section below), followed by the visual presentation of the concept. For each concept, the distractor will be identified by selecting that with the average score of figurative representativeness within the pool of distractors, based on results from the previously performed survey (see above). By pressing two separate buttons on the keyboard indicating different option responses—i.e., “related” vs. “unrelated”—participants will decide as fast as possible whether the previously delivered image represents the concept. The association between buttons and option responses will be counterbalanced between participants (Figure 2, Panel B).

**Figure 2. Panel A)** An example of the trial structure in the Visual Search Task for the concept of “Truce”. **Panel B)** An example of the trial in the Semantic Judgment Task for the concept of “Trench”.



Notably, both tasks will be alternatively performed in two different within-subjects experimental conditions:

1) Verbal interference condition: Before performing the tasks, participants will be asked to read two pseudowords— ten pseudowords were generated with the online free software: [http://www.lexique.org/?page\\_id=582](http://www.lexique.org/?page_id=582); they were chosen randomly from a pool of options among those offered by the software and composed of trigrams, with a maximum length of six letters. Participants were asked to keep them in memory because, every eight trials, they will be asked to report them in the correct order of appearance.

2) Visual interference condition: Participants will be asked to observe two non-canonical geometrical shapes selected from Canva and keep them in memory. Every eight trials, they will be asked to choose, among others, the presented geometrical shapes in the correct order of appearance. These two conditions were inserted to control the adoption of compensatory verbal strategies by aphantasic individuals, who might easily accomplish the task without relying on mental imagery resources.

Each task and experimental condition will be counterbalanced between subjects, and concrete/abstract descriptions will be randomly administered trial-by-trial.

At the end of the behavioral tasks, participants will be asked to evaluate target concepts on some semantic dimensions, such as Perceptual strength, i.e., the extent to which each concept activates sensory modality (hearing, smell, taste, touch, vision; PS - Connell & Lynott, 2012; 2014; Lynott &

Connell, 2013; Villani et al., 2019); Imageability, referring to the level of imageability of entities to which the concept refers (IMG - Paivio et al., 1988; Paivio, 1990; Villani et al., 2019); Concreteness~Abstractness, relying on the amount of abstractness of a concept (ABS - Paivio et al., 1988; Paivio, 1990; Villani et al., 2019; Falcinelli et al., 2023); and Interoception, indicating to which extent a concept activate inner processes (INT - Connell et al., 2018; Villani et al., 2019; 2021). Participants will also be asked to provide socio-demographic information (e.g., sex, age).

### **Pilot Timings**

A pilot study was conducted on ten participants to establish suitable timings for performing our study's visual search and semantic judgment tasks. The pilot employed a within-subjects design to ensure consistency across tasks.

Before the pilot, initial timings for the visual search task were drawn from the study by Zou et al. (2024), where they found that, on average, it took 3 seconds to process each image in a complex scene. Given that our task presented five scenes in an array, the total time allocated for the pilot version of the visual search task was 15,000 milliseconds. Initial timings for the semantic judgment task were adopted from Azimi et al. (2020), who implemented a simultaneous display of lexical stimuli for 5,000 ms judgments. We used this timing in the pilot since our semantic task requires a simultaneous display of a concept and an un/related image.

Following the pilot study, adjustments were made to the timings based on participants' performance both in terms of accuracy and reaction times (RTs). The timing for the semantic judgment task was reduced to 1.2 seconds and for the visual search task to 4.1 seconds per trial (according to the stated above criteria). These timings are estimated considering the round-up of one-third added to the slower RTs values in case of correct responses in each task. These adjustments are aimed to ensure that the timings did not facilitate stimulus learning while providing sufficient time to avoid excessive missing responses. These considerations are particularly relevant for the population of aphantasics, whose performance we will investigate and expect to be slower than controls.

### **Specific Hypotheses**

#### *Semantic Relatedness Judgement Task:*

Verbal interference condition:

1. Aphantasic individuals might show a worse performance (i.e., slower RTs and worse accuracy) when associating (especially) concrete concepts with the corresponding images as compared with non-aphantasic individuals;

2. Non-aphantasic individuals might show the typical Concreteness Effect—displaying faster RTs and better accuracy when associating concrete concepts rather than abstract ones with the corresponding images.

Visual interference condition:

1. Aphantasic individuals might show a worse performance (i.e., slower RTs and worse accuracy) when associating concrete concepts with the corresponding images compared to non-aphantasic individuals.
2. Non-aphantasic individuals might show a worse performance (i.e., slower RTs and worse accuracy) when associating concrete concepts with the corresponding images compared to the verbal interference condition.

*Visual Search Task:*

Verbal interference condition:

1. Aphantasic individuals might show slower RTs and less accuracy when identifying the target image among the four distractors for abstract and concrete concepts than non-aphantasic individuals.
2. Non-aphantasic individuals might show better accuracy and faster RTs in identifying the target image representing concrete concepts than abstract ones.

Visual interference condition:

1. Aphantasic individuals might perform similarly (in RTs and accuracy) when associating concrete and abstract concepts with the corresponding images.
2. Compared to the verbal interference condition, the non-aphantasic individuals might show a slightly reduced accuracy and be slower in RTs when associating concrete and abstract concepts with the corresponding images.

## **Sample**

An independent sample of 30 Aphantasic and 30 non-Aphantasic individuals, all English native speakers, will be tested in the current study. This sample size has been set after performing a power analysis through MorePower 3 software (Faul et al., 2007). Specifically, we estimated the sample

size to achieve a large effect size ( $\eta^2 = .12$ ), with a power of .80%, an alpha error of .05, and a correlation among repeated measures of .50 for an ANOVA with a within factor of two levels Category of Word: abstract, concrete, another within factor of two levels (Type of Interference: verbal interference condition, visual interference condition), a between factor of two levels (Group: Aphantasic group, Non-aphantasic group) and their interaction.

#### *Manipulated variables*

- Group: Aphantasic group, Non-aphantasic group;
  - Category of Word: Abstract, Concrete;
  - Type of Interference: Visual Interference condition, Verbal interference condition.

#### *Measured variables*

For both tasks:

- Accuracy in responses
- Reaction times (in ms)

### **A2.3. Analysis plan**

#### **Visual search task**

##### *Accuracy analysis*

We will employ Generalized Linear Mixed models (GLMMs) with a binomial family. Accuracy scores will be entered as the dependent variable, Group (Aphantasic, Non-Aphantasic), Category of Word (Abstract, Concrete), Type of Interference (Visual Interference, Verbal Interference) and their interaction as fixed factors, and Participants and Words as random factors.

##### *RTs analysis*

We will employ Linear Mixed models (LMMs). The model will fit reaction times as a dependent variable, Group (Aphantasic, Non-Aphantasic), Category of Word (Abstract, Concrete), and Type of Interference (Visual Interference, Verbal Interference) and their interaction as fixed factors, and Participants and Words as random intercepts. In case the response variable is not normally distributed we will revert to GLMMs".

## **Semantic relatedness judgment task**

### *Accuracy analysis*

We will employ Generalized Linear Mixed models (GLMMs) with a binomial family. Accuracy scores will be entered as the dependent variable, Group (Aphantasic, No Aphantasic), Category of Word (Abstract, Concrete), Type of Interference (Visual Interference, Verbal Interference) and their interaction as fixed factors, and Participants and Words as random factors.

### *RTs analysis*

We will employ Linear Mixed models (LMMs). The model will fit reaction times as a dependent variable, Group (Aphantasic, Non-Aphantasic), Category of Word (Abstract, Concrete), and Type of Interference (Visual Interference, Verbal Interference) and their interaction as fixed factors, and Participants and Words as random intercepts. In case the response variable is not normally distributed we will revert to GLMMs.

## References

- Abdulkarim, Y. I., Deng, L., Luo, H., Huang, S., Karaaslan, M., Altıntaş, O., ... & Al-badri, K. S. L. (2020). Design and study of a metamaterial based sensor for the application of liquid chemicals detection. *Journal of Materials Research and Technology*, 9(5), 10291-10304.
- Abdulkarim, Z., & Ehrsson, H. H. (2016). No causal link between changes in hand position sense and feeling of limb ownership in the rubber hand illusion. *Attention, Perception, & Psychophysics*, 78, 707-720.
- Aglioti, S., DeSouza, J. F., and Goodale, M. A. (1995). Size-contrast illusions deceive the eye but not the hand. *Curr. Biol.* 5, 679–685.
- Aimola Davies, A. M., White, R. C., & Davies, M. (2013). Spatial limits on the nonvisual self-touch illusion and the visual rubber hand illusion: Subjective experience of the illusion and proprioceptive drift. *Consciousness and Cognition*, 22(2), 613–636.
- Albers, C., Schmiedt, J. T., & Pawelzik, K. R. (2013). Theta-specific susceptibility in a model of adaptive synaptic plasticity. *Frontiers in computational neuroscience*, 7, 170.
- Amit, E., & Greene, J. D. (2012). You see, the ends don't justify the means: Visual imagery and moral judgment. *Psychological science*, 23(8), 861–868.
- Amodio, D. M. (2014). The neuroscience of prejudice and stereotyping. *Nature Reviews Neuroscience*, 15(10), 670-682.
- Amodio, D. M., & Devine, P. G. (2006). Stereotyping and evaluation in implicit race bias: evidence for independent constructs and unique effects on behavior. *Journal of personality and social psychology*, 91(4), 652.
- Amodio, D. M., & Frith, C. D. (2016). Meeting of minds: the medial frontal cortex and social cognition. *Discovering the social mind*, 183-207.
- Amodio, D. M., Devine, P. G., & Harmon-Jones, E. (2008). Individual differences in the regulation of intergroup bias: the role of conflict monitoring and neural signals for control. *Journal of personality and social psychology*, 94(1), 60.
- Amoruso, L., Finisguerra, A., & Urgesi, C. (2018). Contextualizing action observation in the predictive brain: Causal contributions of prefrontal and middle temporal areas. *Neuroimage*, 177, 68-78.
- Andrade, J., May, J., Deeprase, C., Baugh, S. J., & Ganis, G. (2014). Assessing vividness of mental imagery: The Plymouth sensory imagery questionnaire. *British Journal of Psychology*, 105(4), 547–563.
- Andrade, L. H., Alonso, J., Mneimneh, Z., Wells, J. E., Al-Hamzawi, A., Borges, G., Bromet, E., Bruffaerts, R., de Girolamo, G., de Graaf, R., Florescu, S., Gureje, O., Hinkov, H. R., Hu, C., Huang, Y., Hwang, I., Jin, R., Karam, E. G., Kovess-Masfety, V., ... Kessler, R. C. (2014). Barriers to mental health treatment: Results from the WHO World Mental Health surveys. *Psychological Medicine*, 44(6), 1303–1317.
- Arbib, M. A., Liebal, K., & Pika, S. (2008). Primate vocalization, gesture, and the evolution of human language. *Current anthropology*, 49(6), 1053-1076.
- Arbib, M., Bonaiuto, J., Jacobs, S., and Frey, S. (2009). Tool use and the distalization of the end-effector. *Psychol. Res.* 73, 441–462.
- Argyle, M., & Dean, J. (1965). Eye-contact, distance and affiliation. *Sociometry*, 28(3), 289–304.
- Arntz, A., Tiesema, M., & Kindt, M. (2007). Treatment of PTSD: A comparison of imaginal exposure with and without imagery rescripting. *Journal of behavior therapy and experimental psychiatry*, 38(4), 345-370.
- Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, 63(4), 596–612.
- Aron, E. N., & Aron, A. (1997). Sensory-processing sensitivity and its relation to introversion and emotionality. *Journal of Personality and Social Psychology*, 73(2), 345–368. <https://doi.org/10.1037/0022-3514.73.2.345>
- Auguie B (2017). `_gridExtra`: Miscellaneous Functions for "Grid" Graphics. R package version 2.3, [https://CRAN.R-project.org/package=\\_gridExtra](https://CRAN.R-project.org/package=_gridExtra)
- Atmaca, S., Sebanz, N., & Knoblich, G. (2011). The joint flanker effect: Sharing tasks with real and imagined co-actors. *Experimental brain research*, 211, 371-385.
- Avenanti, A., Candidi, M., & Urgesi, C. (2013). Vicarious motor activation during action perception: beyond correlational evidence. *Frontiers in human neuroscience*, 7, 185.
- Avenanti, A., Sirigu, A., & Aglioti, S. M. (2010). Racial bias reduces empathic sensorimotor resonance with other-race pain. *Current biology*, 20(11), 1018-1022.
- Avillac, M., Hamed, S. B., & Duhamel, J. R. (2007). Multisensory integration in the ventral intraparietal area of the macaque monkey. *Journal of Neuroscience*, 27(8), 1922-1932.

- Azevedo, R. T., Macaluso, E., Avenanti, A., Santangelo, V., Cazzato, V., & Aglioti, S. M. (2013). Their pain is not our pain: brain and autonomic correlates of empathic resonance with the pain of same and different race individuals. *Human brain mapping*, 34(12), 3168-3181.
- Azimi, T., Qoreishi, Z. S., Nilipour, R., & Farazi, M. (2020). Developing a Semantic Similarity Judgment Test for Persian Action Verbs and Non-action Nouns in Patients With Brain Injury and Determining its Content Validity. *Archives of Rehabilitation*, 21(2), 154-181.
- Banks, B., Borghi, A. M., Fagier, R., Fini, C., Jonauskaitė, D., Mazzuca, C., Montalti, M., Villani, C., & Woodin, G. (2022). Consensus paper: Current perspectives on abstract concepts and future research directions. (in press)
- Barca, L., Burani, C., & Arduino, L. S. (2002). Word naming times and psycholinguistic norms for Italian nouns. *Behavior research methods, instruments, & computers*, 34, 424-434.
- Barca, L., Mazzuca, C., & Borghi, A. M. (2017) Pacifier overuse and conceptual relations of abstract and emotional concepts. *Frontiers in Psychology*, 8, 2014.
- Barca, L., Mazzuca, C., & Borghi, A. M. (2020) Overusing the pacifier during infancy sets a foot print on abstract words processing. *Journal of Child Language*, 47, 1084-1099.
- Barros-Loscertales, A., González, J., Pulvermüller, F., Ventura-Campos, N., Bustamante, J. C., Costumero, V., ... & Ávila, C. (2012). Reading salt activates gustatory brain regions: fMRI evidence for semantic grounding in a novel sensory modality. *Cerebral cortex*, 22(11), 2554-2563.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and brain sciences*, 22(4), 577-660.
- Barsalou, L. W., & Wiemer-Hastings, K. (2005). Situating Abstract Concepts. In D. Pecher & R. A. Zwaan (Eds.), *Grounding Cognition* (1st ed., pp. 129–163). Cambridge University Press.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *J StatSoftw.* 2015 Oct; 67(1):1–48.
- Batson, C. D., Sager, K., Garst, E., Kang, M., Rubchinsky, K., & Dawson, K. (1997). Is empathy-induced helping due to self–other merging?. *Journal of personality and social psychology*, 73(3), 495.
- Bavelas, J. B. (2022). *Face-to-face dialogue: theory, research, and applications*. Oxford University Press
- Bavelas, J. B., & Gerwing, J. (2011). The listener as addressee in face-to-face dialogue. *International Journal of Listening*, 25(3), 178-198
- Bavelas, J., Gerwing, J., & Healing, S. (2014). Hand and facial gestures in conversational interaction. In T. M. Holtgraves (Ed.), *The Oxford handbook of language and social psychology* (pp. 111–130). Oxford University Press.
- Bazarova, N. N. (2012). Public Intimacy: Disclosure Interpretation and Social Judgments on Facebook. *Journal of Communication*, 62(5), 815–832.
- Becchio, C., Del Giudice, M., Dal Monte, O., Latini-Corazzini, L., and Pia, L. (2013). In your place: neuropsychological evidence for altercentric remapping in embodied perspective taking. *Soc. Cogn. Affect. Neurosci.* 8, 165–170. doi: 10.1093/scan/nsr083
- Bekkering, H., Wohlschläger, A., & Gattis, M. (2000). Imitation of gestures in children is goal-directed. *The Quarterly Journal of Experimental Psychology Section A*, 53(1), 153-164.
- Belardinelli, A., Lohmann, J., Farnè, A., & Butz, M. V. (2018). Mental space maps into the future. *Cognition*, 176, 65–73. <https://doi.org/10.1016/j.cognition.2018.03.007>
- Bender, M. B. (1952). *Disorders in perception; with particular reference to the phenomena of extinction and displacement*. Springfield: Thomas.
- Bender, M.B., Feldman, D.S., (1951). Extinction of taste sensation on double simultaneous stimulation. *Neurology* 2 (3), 195–202
- Bergelson, E., & Swingle, D. (2013). The acquisition of abstract words by young infants. *Cognition*, 127(3), 391-397.
- Bergen, B. (2007). Experimental methods for simulation semantics. *Methods in cognitive linguistics*, pp. 277, 301.
- Berger, C. R., & Calabrese, R. J. (1975). Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. *Human Communication Research*, 1(2), 99–112.
- Bermúdez, J. L., Eilan, N., & Marcel, A. (Eds.). (1998). *The body and the self*. MIT Press.
- Bernasconi, F., Noel, J.P., Park, H., Faivre, N., Seeck, M., Spinelli, L., et al., (2018). Audiotactile and peripersonal space processing around the trunk in human parietal and temporal cortex: an intracranial EEG. *Cereb. Cortex* 28 (9), 3385–3397.
- Berti, A., and Frassinetti, F. (2000). When far becomes near: remapping of space by tool use. *J. Cogn. Neurosci.* 12, 415–420.
- Biermann-Ruben, K., Kessler, K., Jonas, M., Siebner, H. R., Bäumer, T., Münchau, A., et al. (2008). Right hemisphere contributions to imitation tasks. *Eur. J. Neurosci.* 27, 1843–1855.
- Blakemore, S.-J., and Decety, J. (2001). From the perception of action to the understanding of intention. *Nat. Rev. Neurosci.* 2, 561–567.
- Blanke, O. (2012). Multisensory brain mechanisms of bodily self-consciousness. *Nature Reviews Neuroscience*, 13(8), 556-571.

- Blanke, O., Slater, M., & Serino, A. (2015). Behavioral, neural, and computational principles of bodily self-consciousness. *Neuron*, 88(1), 145–166.
- Bockler, A., Knoblich, G., and Sebanz, N. (2011). Giving a helping hand: effects of joint attention on mental rotation of body parts. *Exp. Brain Res.* 211, 531–545. doi: 10.1007/s00221-011-2625-z
- Bonifazi, S, Farnè, A., Rinaldesi, L. and Ladavas, E. (2007). Dynamic size-change of perihand space through tool-use: spatial extension or shift of the multi-sensory area. *Journal of Neuropsychology*, 1: 101-14.
- Borghi, A. M. (2022). Concepts for Which We Need Others More: The Case of Abstract Concepts. *Current Directions in Psychological Science*, 31(3), 238–246.
- Borghi, A. M. (2022). Merging affordances and (abstract) concepts. In *Affordances in Everyday Life: A Multidisciplinary Collection of Essays* (pp. 113–121). Cham: Springer International Publishing.
- Borghi, A. M. (2023). *The freedom of words: abstractness and the power of language*. Cambridge University Press.
- Borghi, A. M. (2023). The freedom of words: Abstractness and the power of language. <https://philpapers.org/rec/BORTFO-12>
- Borghi, A. M., & Binkofski, F. (2014). *Words as social tools: An embodied view on abstract concepts* (Vol. 2). Springer New York.
- Borghi, A. M., & Riggio, L. (2009). Sentence comprehension and simulation of object temporary, canonical and stable affordances. *Brain Research*, 1253, 117-128.
- Borghi, A. M., & Setti, A. (2017). Abstract concepts and aging: An embodied and grounded perspective. *Frontiers in Psychology*, 8, 430.
- Borghi, A. M., & Tummolini, L. (2020). Touch me if you can: The intangible but grounded nature of abstract concepts. *Behavioral and Brain Sciences*, 43, E123.
- Borghi, A. M., & Zarcone, E. (2016). Grounding abstractness: Abstract concepts and the activation of the mouth. *Frontiers in Psychology*, 7, 1498.
- Borghi, A. M., Barca, L., Binkofski, F., & Tummolini, L. (2018a). Varieties of abstract concepts: Development, use and representation in the brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1752), 20170121. <https://doi.org/10.1098/rstb.2017.0121>
- Borghi, A. M., Barca, L., Binkofski, F., & Tummolini, L. (2018b). Abstract concepts, language and sociality: From acquisition to inner speech. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1752), 20170134. <https://doi.org/10.1098/rstb.2017.0134>
- Borghi, A. M., Barca, L., Binkofski, F., Castelfranchi, C., Pezzulo, G., & Tummolini, L. (2019a). Words as social tools: Flexibility, situatedness, language and sociality in abstract concepts: Reply to comments on ‘Words as social tools: Language, sociality and inner grounding in abstract concepts’. *Physics of Life Reviews*, 29, 178–184.
- Borghi, A. M., Barca, L., Binkofski, F., Castelfranchi, C., Pezzulo, G., & Tummolini, L. (2019b). Words as social tools: Language, sociality and inner grounding in abstract concepts. *Physics of Life Reviews*, 29, 120–153.
- Borghi, A. M., Binkofski, F., Castelfranchi, C., Cimatti, F., Scorolli, C., & Tummolini, L. (2017). The challenge of abstract concepts. *Psychological Bulletin*, 143(3), 263.
- Borghi, A. M., Flumini, A., Cimatti, F., Marocco, D., & Scorolli, C. (2011). Manipulating objects and telling words: a study on concrete and abstract words acquisition. *Frontiers in Psychology*, 2:15.
- Borghi, A. M., Osińska, A., Roepstorff, A., & Raczaszek-Leonardi, J. (2023). Concepts in interaction: social engagement and inner experiences. *Philosophical Transactions of the Royal Society B*, 378(1870), 20210351.
- Botvinick, M. (2004). Probing the neural basis of body ownership. *Science*, 305(5685), 782-783.
- Botvinick, M., & Cohen, J. (1998). Rubber hands ‘feel’ touch that eyes see. *Nature*, 391(6669), 756-756.
- Boukarras, S., Era, V., Aglioti, S. M., & Candidi, M. (2020). Modulation of preference for abstract stimuli following competence-based social status primes. *Experimental Brain Research*, 238(1), 193-204.
- Boukarras, S., Era, V., Aglioti, S. M., & Candidi, M. (2021). Competence-based social status and implicit preference modulate the ability to coordinate during a joint grasping task. *Scientific reports*, 11(1), 5321.
- Bradley, M. M., & Lang, P. J. (1999). Affective norms for English words (ANEW): Instruction manual and affective ratings (Vol. 30, No. 1, pp. 25–36). Technical report C-1, The center for research in psychophysiology, University of Florida.
- Bradner, E., & Mark, G. (2002). Why distance matters: Effects on cooperation, persuasion and deception. *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work*, 226–235. <https://doi.org/10.1145/587078.587110>
- Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80, 1–28. <https://doi.org/10.18637/jss.v080.i01>
- Brass, M., Derrfuss, J., & von Cramon, D. Y. (2005). The inhibition of imitative and overlearned responses: a functional double dissociation. *Neuropsychologia*, 43(1), 89-98.
- Brass, M., Derrfuss, J., Matthes-von Cramon, G., & von Cramon, D. Y. (2003). Imitative response tendencies in patients with frontal brain lesions. *Neuropsychology*, 17(2), 265.

- Bremmer, F., Schlack, A., Shah, N. J., Zafiris, O., Kubischik, M., Hoffmann, K. P., ... & Fink, G. R. (2001). Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalencies between humans and monkeys. *Neuron*, 29(1), 287-296.
- Briggs, G. G., & Nebes, R. D. (1975). Patterns of hand preference in a student population. *Cortex*, 11(3), 230-238.
- Brozzoli, C., Cardinali, L., Pavani, F., & Farnè, A. (2010). Action-specific remapping of peripersonal space. *Neuropsychologia*, 48(3), 796-802.
- Brozzoli, C., Demattè, M. L., Pavani, F., Frassinetti, F., & Farnè, A. (2006). Neglect and extinction: within and between sensory modalities. *Restorative neurology and neuroscience*, 24(4-6), 217-232.
- Brozzoli, C., Ehrsson, H. H., & Farnè, A. (2014). Multisensory representation of the space near the hand: from perception to action and interindividual interactions. *The Neuroscientist*, 20(2), 122-135.
- Brozzoli, C., Gentile, G., Bergouignan, L., & Ehrsson, H. H. (2013). A shared representation of the space near oneself and others in the human premotor cortex. *Current Biology*, 23(18), 1764-1768.
- Brozzoli, C., Makin, T., Cardinali, L., Holmes, N., Farnè, A., (2012). Peripersonal space: a multisensory interface for body-object interaction. In: Micah, M.M., Wallace, M.T. (Eds.), *The Neural Bases of Multisensory Processes*, pp. 449-466 Boca Raton (FL).
- Brozzoli, C., Pavani, F., Urquizar, C., Cardinali, L., & Farnè, A. (2009). Grasping actions remap peripersonal space. *Neuroreport*, 20(10), 913-917.
- Brozzoli, Claudio, Makin, Tamar R., Cardinali, Lucilla, Holmes, Nicholas P. and Farne, Alessandro (2011). Peripersonal space: a multisensory interface for body-object interactions. In: Murray, M. M. and Wallace, M. T. (eds.) *The neural bases of multisensory processes*. Frontiers in Neuroscience.
- Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior research methods*, 46, 904-911.
- Buccino, G., Sato, M., Cattaneo, L., Rodà, F., & Riggio, L. (2009). Broken affordances, broken objects: a TMS study. *Neuropsychologia*, 47(14), 3074-3078.
- Bürkner, P.-C., & Vuorre, M. (2019). Ordinal regression models in psychology: A tutorial. *Advances in Methods and Practices in Psychological Science*, 2(1), 77-101.
- Butterfill, S. (2012). Joint action and development. *The Philosophical Quarterly*, 62(246), 23-47.
- Bzdok, D., & Dunbar, R. I. M. (2020). The Neurobiology of Social Distance. *Trends in Cognitive Sciences*, 24(9), 717-733.
- Caggiano, V., Fogassi, L., Rizzolatti, G., Thier, P., and Casile, A. (2009). Mirror neurons differentially encode the peripersonal and extrapersonal space of monkeys. *Science* 324, 403-406.
- Caiani, SZ. (2014). Extending the notion of affordance. *Phenomenology and the Cognitive Sciences*, 13, 275-293.
- Cairncross, F. (2001). *The Death of Distance: How the Communications Revolution is Changing Our Lives*. Harvard Business School Press
- Campbell, J. I. D., & Thompson, V. A. (2012). MorePower 6.0 for ANOVA with relational confidence intervals and Bayesian analysis. *Behavior Research Methods*, 44(4), 1255-1256
- Candidi, M., Curioni, A., Donnarumma, F., Sacheli, L. M., & Pezzulo, G. (2015). Interactional leader-follower sensorimotor communication strategies during repetitive joint actions. *Journal of The Royal Society Interface*, 12(110), 20150644. <https://doi.org/10.1098/rsif.2015.0644>
- Candidi, M., Sacheli, L. M., Era, V., Canzano, L., Tieri, G., & Aglioti, S. M. (2017). Come together: human-avatar on-line interactions boost joint-action performance in apraxic patients. *Social Cognitive and Affective Neuroscience*, 12(11), 1793-1802. <https://doi.org/10.1093/scan/nsx114>
- Candini, M., Giuberti, V., Manattini, A., Grittani, S., di Pellegrino, G., & Frassinetti, F. (2017). Personal space regulation in childhood autism: Effects of social interaction and person's perspective. *Autism Research: Official Journal of the International Society for Autism Research*, 10(1), 144-154.
- Canzoneri E, Magosso E, Serino A. (2012). Dynamic sounds capture the boundaries of peripersonal space representation in humans. *PLoS One*;7(9):e44306.
- Cao, Y., Summerfield, C., Park, H., Giordano, B. L., & Kayser, C. (2019). Causal inference in the multisensory brain. *Neuron*, 102(5), 1076-1087.e8. <https://doi.org/10.1016/j.neuron.2019.03.043>
- Caramazza, A., & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: The animate-inanimate distinction. *Journal of cognitive neuroscience*, 10(1), 1-34.
- Cardinali, L., Frassinetti, F., Brozzoli, C., Urquizar, C., Roy, A. C., & Farnè, A. (2009). Tool-use induces morphological updating of the body schema. *Current biology*, 19(12), R478-R479.
- Cardini, F., Bertini, C., Serino, A., & Ladavas, E. (2012). Emotional modulation of visual remapping of touch. *Emotion*, 12(5), 980.
- Cartaud, A., Ott, L., Iachini, T., Honoré, J., and Coello, Y. (2020). The influence of facial expression at perceptual threshold on electrodermal activity and social comfort distance. *Psychophysiology* 57: e13600.

- Cartaud, A., Ruggiero, G., Ott, L., Iachini, T., and Coello, Y. (2018). Physiological response to facial expressions in peripersonal space determines interpersonal distance in a social interaction context. *Front. Psychol.* 9:657.
- Ceunen, E., Vlaeyen, J. W., Van Diest, I. (2016). On the Origin of Interoception. *Front. Psychol.* 7, 743.
- Chaminade, T., Meltzoff, A. N., & Decety, J. (2005) An fMRI study of imitation: action representation and body schema. *Neuropsychologia* 43, 115–127.
- Chang, J. H., Hodgins, H. S., Lin, Y. C., & Huang, C. L. (2014). On the flexibility of the automatic Us-Them divide. In N. Weinstein (Ed.), *Human motivation and interpersonal relationships: Theory, research, and applications* (pp. 97–119). Dordrecht: Springer.
- Chemero, A. (2003). Information for perception and information processing. *Minds and Machines*, 13, 577-588.
- Chouikhi, H., Charrad, M., & Ghazzali, N. (2015). A comparison study of clustering validity indices. In 2015 global summit on computer & information technology (GSCIT), *IEEE* pp. 1–4.
- Christensen, M. S., Lundbye-Jensen, J., Geertsen, S. S., Petersen, T. H., Paulson, O. B., & Nielsen, J. B. (2007). Premotor cortex modulates somatosensory cortex during voluntary movements without proprioceptive feedback. *Nature Neuroscience*, 10(4), 417-419.
- Clark, A. (1997). *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press.
- Clark, H. H. (1996). *Using language* (pp. xi, 432). Cambridge University Press.
- Cléry, J., & Hamed, S. B. (2018). Frontier of self and impact prediction. *Frontiers in Psychology*, 9, 1073.
- Cohen, J. (2013). Statistical power analysis for the behavioral sciences. *Routledge*.
- Colby, C. L., & Duhamel, J. R. (1996). Spatial representations for action in parietal cortex. *Cognitive Brain Research*, 5(1-2), 105-115.
- Colby, C. L., and Goldberg, M. E. (1999). Space and attention in parietal cortex. *Annu. Rev. Neurosci.* 22, 319–349.
- Connell, L., & Lynott, D. (2012). Strength of perceptual experience predicts word processing performance better than concreteness or imageability. *Cognition*, 125(3), 452–465.
- Connell, L., & Lynott, D. (2012). Strength of perceptual experience predicts word processing performance better than concreteness or imageability. *Cognition*, 125(3), 452–465.
- Connell, L., & Lynott, D. (2014). Principles of representation: Why you can't represent the same concept twice. *Topics in Cognitive Science*, 6(3), 390–406.
- Connell, L., Lynott, D., & Banks, B. (2018). Interoception: the forgotten modality in the perceptual grounding of abstract and concrete concepts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1752), 20170143.
- Connor, P., Weeks, M., Glaser, J., Chen, S., & Keltner, D. (2023). Intersectional implicit bias: Evidence for asymmetrically compounding bias and the predominance of target gender. *Journal of Personality and Social Psychology*, 124(1), 22.
- Constable, M. D., Kritikos, A., & Bayliss, A. P. (2011). Grasping the concept of personal property. *Cognition*, 119(3), 430-437.
- Contreras-Huerta, L. S., Baker, K. S., Reynolds, K. J., Batalha, L., & Cunnington, R. (2013). Racial bias in neural empathic responses to pain. *PloS one*, 8(12), e84001.
- Contreras-Huerta, L. S., Hielscher, E., Sherwell, C. S., Rens, N., & Cunnington, R. (2014). Intergroup relationships do not reduce racial bias in empathic neural responses to pain. *Neuropsychologia*, 64, 263-270.
- Cooke, D. F., & Graziano, M. S. A. (2003). Defensive movements evoked by air puff in monkeys. *Journal of Neurophysiology*, 90(1), 3317-3329.
- Cooley, E., Lei, R. F., & Ellerkamp, T. (2018). The mixed outcomes of taking ownership for implicit racial biases. *Personality and Social Psychology Bulletin*, 44(10), 1424-1434.
- Corballis, M. C. (2010). Mirror neurons and the evolution of language. *Brain and language*, 112(1), 25-35.
- Cosmides, L., Tooby, J., & Kurzban, R. (2003). Perceptions of race. *Trends in cognitive sciences*, 7(4), 173-179.
- Costall, A. (1995). Socializing affordances. *Theory & Psychology*, 5(4), 467–481
- Craig, A. D. (2010). The sentient self. *Brain Struct. Funct.* 214, 563–577. <http://dx>.
- Craighero, L., Fadiga, L., Rizzolatti, G., and Umiltà C. A. (1999). Movement for perception: a motor-visual attentional effect. *Journal of Experimental Psychology*, 25/6:1673-92.
- Crutcher, M. D., & DeLong, M. R. (1984). Single cell studies of the primate putamen: II. Relations to direction of movement and pattern of muscular activity. *Experimental Brain Research*, 53, 244-258.
- Curioni, A., Knoblich, G. K., Sebanz, N., & Sacheli, L. M. (2020). The engaging nature of interactive gestures. *Plos one*, 15(4), e0232128.
- Curioni, A., Minio-Paluello, I., Sacheli, L. M., Candidi, M., & Aglioti, S. M. (2017). Autistic traits affect interpersonal motor coordination by modulating strategic use of role-based behavior. *Molecular Autism*, 8(1).
- Daft, R., & Lengel, R. (1983). Information Richness. A New Approach to Managerial Behavior and Organization Design. *Research in Organizational Behavior*, 6, 73
- Davare, M., Kraskov, A., Rothwell, J. C. & Lemon, R. N. (2011). Interactions between areas of the cortical grasping network. *Curr. Opin. Neurobiol.* 21, 565–570

- Davis, C. P., Altmann, G. T., & Yee, E. (2020). Situational systematicity: A role for schema in understanding the differences between abstract and concrete concepts. *Cognitive Neuropsychology*, 37, 142–153.
- Davis, M. H. (1980). Interpersonal reactivity index. APA PsycTests. <https://doi.org/10.1037/t01093-000>
- De Waal, F. B., & Preston, S. D. (2017). Mammalian empathy: behavioural manifestations and neural basis. *Nature Reviews Neuroscience*, 18(8), 498–509.
- De' Sperati, C., and Stucchi, D. (1997). Recognizing the motion of a graspable object is guided by handedness. *Neuroreport*, 8: 2761–2765.
- Deacon, T. W. (1998). *The symbolic species: The co-evolution of language and the brain* (Vol. 202). WW Norton & Company.
- Decety, J., Michalska, K. J., Akitsuki, Y., & Lahey, B. B. (2009). Atypical empathic responses in adolescents with aggressive conduct disorder: a functional MRI investigation. *Biological psychology*, 80(2), 203–211.
- Dehaene, S. (1997). *The Number Sense*; Oxford University Press: New York, NY, USA
- Della Rosa, P. A., Catricalà, E., Vigliocco, G., & Cappa, S. F. (2010). Beyond the abstract—concretedichotomy: Mode of acquisition, concreteness, imageability, familiarity, age of acquisition, context availability, and abstractness norms for a set of 417 Italian words. *Behavior Research Methods*, 42(4), 1042–1048.
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>
- Desai, R. H., Reilly, M., & van Dam, W. (2018). The multifaceted abstract brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1752), 20170122.
- Désy, M. C., & Théoret, H. (2007). Modulation of motor cortex excitability by physical similarity with an observed hand action. *PLoS One*, 2(10), e971.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of personality and social psychology*, 56(1), 5.
- di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, 91, 176–180.
- di Pellegrino, G., Ladavas, E., & Farné, A. (1997). Seeing where your hands are. *Nature*, 388(6644), 730–730.
- di Pellegrino, G., Rafal, R., & Tipper, S. P. (2005). Implicitly evoked actions modulate visual selection: Evidence from parietal extinction. *Current Biology*, 15(16), 1469–1472.
- Diveica, V., Pexman, P. M., & Binney, R. J. (2022). Quantifying social semantics: An inclusive definition of socialness and ratings for 8388 English words. *Behavior Research Methods*, 1–13.
- Do Bú, E. A., Madeira, F., Pereira, C. R., Hagiwara, N., & Vala, J. (2023). Intergroup time bias and aversive racism in the medical context. *Journal of Personality and Social Psychology*.
- Dove, G. (2009). Beyond perceptual symbols: A call for representational pluralism. *Cognition*, 110(3), 412–431.
- Dove, G. (2011). On the need for embodied and dis-embodied cognition. *Frontiers in Psychology*, 1, 242.
- Dove, G. (2014). Thinking in words: Language as an embodied medium of thought. *Topics in Cognitive Science*, 6(3), 371–389.
- Dove, G. (2021). The challenges of abstract concepts. *Handbook of Embodied Psychology: Thinking, Feeling, and Acting*, 171–195
- Dove, G. (2022). *Abstract concepts and the embodied mind: Rethinking grounded cognition*. Oxford University Press
- Dovidio, J. F., Kawakami, K., Gaertner, S. L., ten Vergert, M., Hodson, G., & Falvo, R. (2005). Understanding and addressing racial bias: Insights from Allport. In *Social psychology of prejudice: historical and contemporary issues*/edited by Christian S. Crandall, Mark Schaller. Lawrence, Kan.: Lewinian Press, c2005.. Lawrence, Kan.: Lewinian Press.
- Dreyer, F. R., & Pulvermüller, F. (2018). Abstract semantics in the motor system?—An event-related fMRI study on passive reading of semantic word categories carrying abstract emotional and mental meaning. *Cortex*, 100, 52–70.
- Duhamel, J.R., Colby, C.L., and Goldberg, M.E. (1998). Ventral Intraparietal area of the macaque: congruent visual and somatic response properties. *Journal of Neurophysiology*, 79: 126–36.
- Dunbar, R. I. (1998). The social brain hypothesis. *Evolutionary Anthropology: Issues, News, and Reviews*, 6(5), 178–190.
- Durgin, F. H., Evans, L., Dunphy, N., Klostermann, S., & Simmons, K. (2007). Rubber hands feel the touch of light. *Psychological Science*, 18(2), 152–157.
- Earls, H. A., Englander, Z. A., & Morris, J. P. (2013). Perception of race-related features modulates neural activity associated with action observation and imitation. *Neuroreport*, 24(8), 410–413.
- Ehrsson, H. H. (2012). The concept of body ownership and its relation to multisensory integration. In B. E. Stein (Ed), *The new handbook of multisensory processes* (pp. 775–792). MIT Press
- Ehrsson, H. H. (2020). Multisensory processes in body ownership. *Multisensory perception*, 179–200.
- Ehrsson, H. H., Holmes, N. P., & Passingham, R. E. (2005). Touching a rubber hand: feeling of body ownership is associated with activity in multisensory brain areas. *Journal of neuroscience*, 25(45), 10564–10573.

- Ellison, N. B., Gray, R., Lampe, C., & Fiore, A. T. (2014). Social capital and resource requests on Facebook. *New Media & Society*, 16(7), 1104–1121.
- Engel, A. K. (2010). “Directive minds: how dynamics shapes cognition,” in *Enaction: Towards a New Paradigm for Cognitive Science*, eds J. Stewart, O. Gapenne, and E. Di Paolo (Cambridge, MA: MIT Press), 219–243
- Engel, A. K., Maye, A., Kurthen, M., & König, P. (2013). Intrinsic coupling modes: multiscale interactions in ongoing brain activity. *Neuron* 80, 867–886.
- Era, V., Aglioti, S. M., Mancusi, C., & Candidi, M. (2020). Visuo-motor interference with a virtual partner is equally present in cooperative and competitive interactions. *Psychological Research*, 84(3), 810–822.
- Era, V., Candidi, M., Gandolfo, M., Sacheli, L. M., & Aglioti, S. M. (2018). Inhibition of left anterior intraparietal sulcus shows that mutual adjustment marks dyadic joint-actions in humans. *Social Cognitive and Affective Neuroscience*, 13(5), 492-500.
- Era, V., Candidi, M., Pezzetta, R., Pulcini, C., D’Antonio, S., Zabberoni, S., ... & Aglioti, S. M. (2023). The dopaminergic system supports flexible and rewarding dyadic motor interactive behaviour in Parkinson’s Disease. *Social Cognitive and Affective Neuroscience*, 18(1), nsac040.
- Era, V., Carnevali, L., Thayer, J. F., Candidi, M., & Ottaviani, C. (2021). Dissociating cognitive, behavioral and physiological stress-related responses through dorsolateral prefrontal cortex inhibition. *Psychoneuroendocrinology*, 124, 105070.
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870), 429–433. <https://doi.org/10.1038/415429a>
- Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (2000). Visuomotor neurons: Ambiguity of the discharge or ‘motor’ perception?. *International journal of psychophysiology*, 35(2-3), 165-177.
- Falcinelli, I., Fini, C., Mazzuca, C., & Borghi, A. M. (2023). The TECo Database: Ecological and Technological as concepts at the interface between abstractness and concreteness.
- Falcinelli, I., Fini, C., Mazzuca, C., & Borghi, A. M. (2023, August 1). The TECo Database: Ecological and Technological as concepts at the interface between abstractness and concreteness. <https://doi.org/10.31234/osf.io/vcn8u>
- Farmer, H., & Maister, L. (2017). Putting ourselves in another’s skin: using the plasticity of self-perception to enhance empathy and decrease prejudice. *Social Justice Research*, 30, 323-354.
- Farmer, H., & Tsakiris, M. (2012). The bodily social self: A link between phenomenal and narrative selfhood. *Review of Philosophy and Psychology*, 3, 125-144.
- Farnè, A. and Ladavas, E. (2000). Dynamic size-change of hand peripersonal space following tool use. *NeuroReport*, 11: 1645-9.
- Farnè, A., Demattè, M. L., & Ladavas, E. (2005a). Neuropsychological evidence of modular organization of the near peripersonal space. *Neurology*, 65(11), 1754-1758.
- Farnè, A., Iriki, A., and Ladavas, E. (2005b). Shaping multisensory action-space with tools: evidence from patients with cross-modal extinction. *Neuropsychologia*, 43: 238-48.
- Farnè, A., Pavani, F., Meneghello, F., & Ladavas, E. (2000). Left tactile extinction following visual stimulation of a rubber hand. *Brain*, 123(11), 2350–2360.
- Farnè, A., Serino, A., & Ladavas, E. (2007). Dynamic size-change of peri-hand space following tool-use: determinants and spatial characteristics revealed through cross-modal extinction. *Cortex*, 43(3), 436-443.
- Ferraina, S., Battaglia-Mayer, A., Genovesio, A., Archambault, P., and Caminiti, R. (2009). Parietal encoding of action in depth. *Neuropsychologia* 47, 1409–1420.
- Ferrari, P. F., Coudé, G., (2018). Mirror neurons, embodied emotions, and empathy. In K. Z. Meyza E. Knapska (Eds.), *Neuronal correlates of empathy: From rodent to human* (p. 67–77). Elsevier Academic Press.
- Fingerhut, J., & Prinz, J. J. (2018). Wonder, appreciation, and the value of art. *Progress in Brain Research*, 237, 107–128.
- Fini, C., & Borghi, A. M. (2019). Sociality to reach objects and to catch meaning. *Frontiers in Psychology*, 10, 838.
- Fini, C., Bardi, L., Bolis, D., Fusaro, M., Lisi, M. P., Michalland, A. H., & Era, V. (2023a). The social roots of self development: From a bodily to an intellectual interpersonal dialogue. *Psychological Research*, 87(6), 1683–1695.
- Fini, C., Era, V., Da Rold, F., Candidi, M., & Borghi, A. M. (2021). Abstract concepts in interaction: The need of others when guessing abstract concepts smooths dyadic motor interactions. *Royal Society Open Science*, 8(7), 201205.
- Fini, C., Era, V., Da Rold, F., Candidi, M., & Borghi, A. M. (2021). Abstract concepts in interaction: The need of others when guessing abstract concepts smooths dyadic motor interactions. *Royal Society Open Science*, 8(7), 201205.
- Fini, C., Falcinelli, I., Cuomo, G., Era, V., Candidi, M., Tummolini, L., Mazzuca, C., & Borghi, A. M. (2023b). Breaking the ice in a conversation: Abstract words prompt dialogs more easily than concrete ones. *Language and Cognition*, 1–22.
- Fini, C., Verbeke, P., Sieber, S., Moors, A., Brass, M., and Genschow, O. (2020). The influence of threat on perceived spatial distance to out-group members. *Psychol. Res.* 84, 757–764.

- Fini, C., Zannino, G. D., Orsoni, M., Carlesimo, G. A., Benassi, M., & Borghi, A. M. (2022). Articulatory suppression delays processing of abstract words: The role of inner speech. *Quarterly Journal of Experimental Psychology*, 75(7), 1343–1354.
- Foerster, F. R., Borghi, A. M., & Goslin, J. (2020). Labels strengthen motor learning of new tools. *Cortex*, 129, 1-10.
- Fogassi, L. & Luppino, G. (2005). Motor functions of the parietal lobe. *Curr. Opin Neurobiol.* 15, 626–631
- Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science*, 308(5722), 662-667.
- Fogassi, L., Gallese, V., di Pellegrino, G., Fadiga, L., Gentilucci, M., Luppino, G., Matelli, M., Pedotti, A., & Rizzolatti, G. (1992). Space coding by premotor cortex. *Experimental Brain Research*, 89(3), 686–690.
- Fogassi, L., Gallese, V., Fadiga, L., Luppino, G., Matelli, M., & Rizzolatti, G. (1996). Coding of peripersonal space in inferior premotor cortex (area F4). *Journal of neurophysiology*, 76(1), 141-157.
- Fogassi, L., Raos, V., Franchi, G., Gallese, V., Luppino, G., & Matelli, M. (1999). Visual responses in the dorsal premotor area F2 of the macaque monkey. *Experimental brain research*, 128, 194-199.
- Fossataro, C., Galigani, M., Sebastiano, A. R., Bruno, V., Ronga, I., & Garbarini, F. (2023). Spatial proximity to others induces plastic changes in the neural representation of the peripersonal space. *IScience*, 26(1).
- Freeman, J. B., Schiller, D., Rule, N. O., & Ambady, N. (2010). The neural origins of superficial and individuated judgments about ingroup and outgroup members. *Human brain mapping*, 31(1), 150-159.
- Frischen, A., & Tipper, S. P. (2004). Orienting attention via observed gaze shift evokes longer term inhibitory effects: implications for social interactions, attention, and memory. *Journal of Experimental Psychology: General*, 133(4), 516.
- Friston, K. (2011). What is optimal about motor control? *Neuron*, pp. 72, 488–498.
- Friston, K., & Frith, C. (2015). A duet for one. *Consciousness and Cognition*, 36, 390–405.
- Frith, C. D., & Frith, U. (2006). The Neural Basis of Mentalizing. *Neuron*, 4, 531–534.
- Frith, U., & Frith, C. (2010). The social brain: allowing humans to boldly go where no other species has been. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 1537, 165–176.
- Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1431), 459-473.
- Froese, T., Iizuka, H., and Ikegami, T. (2014). Embodied social interaction constitutes social cognition in pairs of humans: a minimalist virtual reality experiment. *Sci. Rep.* 4:3672.
- Fuehrer, E., Voudouris, D., Lezkan, A., Drewing, K., & Fiehler, K. (2022). Tactile suppression stems from specific sensorimotor predictions. *Proceedings of the National Academy of Sciences*, 119(20).
- Fusaro, M., Tieri, G., & Aglioti, S. M. (2019). Influence of cognitive stance and physical perspective on subjective and autonomic reactivity to observed pain and pleasure: An Immersive Virtual Reality Study. *Consciousness and Cognition*, 67, 86–97. <https://doi.org/10.1016/j.concog.2018.11.010>
- Gaesser, B., & Schacter, D. L. (2014). Episodic simulation and episodic memory can increase intentions to help others. *Proceedings of the National Academy of Sciences*, 111(12), 4415-4420.
- Gallace, A., & Spence, C. (2010). The science of interpersonal touch: an overview. *Neuroscience & Biobehavioral Reviews*, 34(2), 246-259.
- Gallagher, M., Colzi, C., & Sedda, A. (2021). Dissociation of proprioceptive drift and feelings of ownership in the somatic rubber hand illusion. *Acta Psychologica*, 212, 103192.
- Gallagher, S., & Meltzoff, A. N. (1996). The earliest sense of self and others: Merleau-Ponty and recent developmental studies. *Philosophical Psychology*, 2, 211–233.
- Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in cognitive sciences*, 8(9), 396-403.
- Gallotti, M., Fairhurst, M.T., Frith, C.D. (2017). Alignment in social interactions. *Consciousness and Cognition*, 48, 253–61.
- Ganel, T., Tanzer, M., and Goodale, M. A. (2008). A double dissociation between action and perception in the context of visual illusions: opposite effects of real and illusory size. *Psychol. Sci.* 19, 221–225.
- Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, 8(1), 8–11.
- Gentilucci, M., Daprati, E., and Gangitano, M. (1998a). Implicit visual analysis in handedness recognition. *Consc. Cogn.* 7:478-493.
- Gentilucci, M., Daprati, E., and Gangitano, M. (1998b). Right-handers and left-handers have different representations of their own hand. *Cognitive Brain Research*, 6: 185-192.
- Gentilucci, M., Fogassi, L., Luppino, G., Matelli, M., Camarda, R., & Rizzolatti, G. (1988). Functional organization of inferior area 6 in the macaque monkey: I. Somatotopy and the control of proximal movements. *Experimental brain research*, 71, 475-490.
- Gessaroli, E., Santelli, E., di Pellegrino, G., & Frassinetti, F. (2013). Personal space regulation in childhood autism spectrum disorders. *PLoS One*, 8(9), e74959.

- Ghaem, O., Mellet, E., Crivello, F., Tzourio, N., Mazoyer, B., Berthoz, A., & Denis, M. (1997). Mental navigation along memorized routes activates the hippocampus, precuneus, and insula. *Neuroreport*, 8(3), 739-744.
- Ghio, M., Vaghi, M. M. S., & Tettamanti, M. (2013). Fine-grained semantic categorization across the abstract and concrete domains. *PLoS One*, 8, e0067090.
- Gianelli, C., Scorolli, C., & Borghi, A. M. (2013). Acting in perspective: the role of body and language as social tools. *Psychological Research*, 77, 40-52.
- Gilbert, S. J., Swencionis, J. K., & Amodio, D. M. (2012). Evaluative vs. trait representation in intergroup social judgments: Distinct roles of anterior temporal lobe and prefrontal cortex. *Neuropsychologia*, 50(14), 3600-3611.
- Gilead, M., Trope, Y., & Liberman, N. (2020). Above and beyond the concrete: The diverse representational substrates of the predictive brain. *Behavioral and Brain Sciences*, 43, e121.
- Gilet, A. L., Mella, N., Studer, J., Grünh, D., & Labouvie-Vief, G. (2013). Assessing dispositional empathy in adults: A French validation of the Interpersonal Reactivity Index (IRI). *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 45(1), 42.
- Goldstein, N. J., & Cialdini, R. B. (2007). The spyglass self: a model of vicarious self-perception. *Journal of personality and social psychology*, 92(3), 402.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends Neurosci.* 15, 20–25.
- Goodale, M. A., Meenan, J. P., Bühlhoff, H. H., Nicolle, D. A., Murphy, K. J., & Racicot, C. I. (1994). Separate neural pathways for the visual analysis of object shape in perception and prehension. *Current Biology*, 4(7), 604-610.
- Goodale, M. A., Milner, A. D., Jakobson, L. S., & Carey, D. P. (1991). A neurological dissociation between perceiving objects and grasping them. *Nature*, 349(6305), 154-156.
- Göschl, F., Engel, A. K., & Fries, U. (2014). Attention modulates Visual-Tactile interaction in spatial pattern matching. *PloS One*, 9(9), e106896. <https://doi.org/10.1371/journal.pone.0106896>
- Grafton, S. T. (2009). Embodied cognition and the simulation of action to understand others. *Ann. N Y Acad. Sci.* 1156, 97–117.
- Grafton, S. T., Fadiga, L., Arbib, M. A., & Rizzolatti, G. (1997). Premotor cortex activation during observation and naming of familiar tools. *Neuroimage*, 6(4), 231-236.
- Granito, C., Scorolli, C., & Borghi A. M. (2015). Naming a Lego World. The Role of Language in the Acquisition of Abstract Concepts. *PLOS ONE*, 10(1):e0114615.
- Graziano, M. S. (1999). Where is my arm? The relative role of vision and proprioception in the neuronal representation of limb position. *Proceedings of the National Academy of Sciences*, 96(18), 10418-10421.
- Graziano, M. S., & Botvinick, M. M. (2002). How the brain represents the body: insights from neurophysiology and psychology. *Common mechanisms in perception and action: Attention and performance XIX*, 19, 136-157.
- Graziano, M. S., & Gross, C. G. (1993). A bimodal map of space: somatosensory receptive fields in the macaque putamen with corresponding visual receptive fields. *Experimental brain research*, 97, 96-109.
- Graziano, M. S., & Gross, C. G. (1995). The representation of extrapersonal space: A possible role for bimodal, visual-tactile neurons. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1021-1034). MIT Press.
- Graziano, M. S., & Gross, C. G. (1996). Multiple pathways for processing visual space. *Attention and performance XVI: Information integration in perception and communication*, 181-207.
- Graziano, M. S., Yap, G. S., and Gross, C. G. (1994). Coding of visual space by premotor neurons. *Science* 266, 1054–1057.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of personality and social psychology*, 74(6), 1464.
- Grèzes, J., Tucker, M., Armony, J., Ellis, R., & Passingham, R. E. (2003). Objects automatically potentiate action: an fMRI study of implicit processing. *European Journal of Neuroscience*, 17(12), 2735-2740.
- Grieve, R., Indian, M., Witteveen, K., Anne Tolan, G., & Marrington, J. (2013). Face-to-face or Facebook: Can social connectedness be derived online? *Computers in Human Behavior*, 29(3),604–609.
- Gross, C. G., & Graziano, M. S. A. (1995). Multiple representations of space in the brain. *Neuroscientist* 1, 43–50.
- Gutsell, J. N., & Inzlicht, M. (2010). Empathy constrained: Prejudice predicts reduced mental simulation of actions during observation of outgroups. *Journal of experimental social psychology*, 46(5), 841-845.
- Hadley, L.V., Novembre, G., Keller, P.E., Pickering, M.J. (2015). Causal role of motor simulation in turn-taking behavior. *Journal of Neuroscience*, 35(50), 16516–20.
- Hall, E. T. (1990). *The Hidden Dimension*. Knopf Doubleday Publishing Group
- Hamilton, A. & Grafton, S. T. (2006). Goal representation in human anterior intraparietal sulcus. *J. Neurosci.* 26, 1133–1137.
- Han, S. (2018). Neurocognitive basis of racial ingroup bias in empathy. *Trends in cognitive sciences*, 22(5), 400-421.
- Hargreaves, D. J. (2012). Musical imagination: Perception and production, beauty and creativity. *Psychology of music*, 40(5), 539–557.

- Harpaintner, M., Trumpp, N. M., & Kiefer, M. (2018). The semantic content of abstract concepts: A property listing study of 296 abstract words. *Frontiers in Psychology*, 9, 1748.
- Harvester Wheatshaeaf Sproull, L., & Kiesler, S. (1986). Reducing Social Context Cues: Electronic Mail in Organizational Communications. *Management Science*, 32(11), 1492–1512
- Hasler, B. S., Spanlang, B., & Slater, M. (2017). Virtual race transformation reverses racial in-group bias. *PloS one*, 12(4), e0174965.
- Hasson, U., Frith, C.D. (2016). Mirroring and beyond: coupled dynamics as a generalized framework for modelling social interactions. *Philosophical Transactions of the Royal Society B*, 371(1693), 20150366.
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional contagion. *Current directions in psychological science*, 2(3), 96-100.
- Hayduk, L. A. (1983). Personal space: where we now stand. *Psychol. Bull.* 94, 293–335. doi: 10.1037/0033-2909.94.2.293
- Head, H., & Holmes, G. (1911). Sensory disturbances from cerebral lesions. *Brain* 34, 102–254.
- Hecht, H., Welsch, R., Viehoff, J., and Longo, M. R. (2019). The shape of personal space. *Acta Psychol.* 193, 113–122.
- Heed, T., Habets, B., Sebanz, N., & Knoblich, G. (2010). Others' actions reduce crossmodal integration in peripersonal space. *Current Biology*, 20(15), 1345-1349.
- Herbert, B., Pollatos, O., (2012). The body in the mind. On the relationship between interoception and embodiment. *Top. Cogn. Sci.* 4, 692–704.
- Hewes, G. W., Andrew, R. J., Carini, L., Choe, H., Gardner, R. A., Kortlandt, A., ... & Wescott, R. W. (1973). Primate communication and the gestural origin of language [and comments and reply]. *Current anthropology*, 14(1/2), 5-24.
- Högman, V., Björkman, M., & Kragic, D. (2013, November). Interactive object classification using sensorimotor contingencies. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 2799-2805). IEEE.
- Holler, J., Bavelas, J., Woods, J., Geiger, M., & Simons, L. (2022). Given-new effects on the duration of gestures and of words in face-to-face dialogue. *Discourse Processes*, 59(8), 619-645
- Holmes, E. A., Arntz, A., & Smucker, M. R. (2007). Imagery rescripting in cognitive behavior therapy: Images, treatment techniques, and outcomes. *Journal of behavior therapy and experimental psychiatry*, 38(4), 297-305.
- Holmes, N. P., & Spence, C. (2004). The body schema and multisensory representation (s) of peripersonal space. *Cognitive processing*, 5, 94-105.
- Holmes, N. P., Calvert, G. A., & Spence, C. (2007). Tool use changes multisensory interactions in seconds: evidence from the crossmodal congruency task. *Experimental Brain Research*, 183, 465-476.
- Holmes, N. P., Sanabria, D., Calvert, G. A., & Spence, C. (2007). Tool-use: capturing multisensory spatial attention or extending multisensory peripersonal space?. *Cortex*, 43(3), 469-489.
- Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social cognition*, 27(6), 949-960.
- Huang, R.S., Chen, C.F., Tran, A.T., Holstein, K.L., Sereno, M.I., (2012). Mapping multisensory parietal face and body areas in humans. *Proc. Natl. Acad. Sci. U. S. A.* 109 (44), 18114–18119.
- Huang, Y. Z., Edwards, M. J., Rounis, E., Bhatia, K. P., & Rothwell, J. C. (2005). Theta burst stimulation of the human motor cortex. *Neuron*, 45(2), 201-206.
- Hyvarinen, J. (1981). Regional distribution of functions in parietal association area 7 of the monkey. *Brain research*, 206(2), 287-303.
- Hyvärinen, J., & Poranen, A. (1974). Function of the parietal associative area 7 as revealed from cellular discharges in alert monkeys. *Brain*, 97(4), 673-692.
- Hyvärinen, J., & Shelepin, Y. E. (1979). Distribution of visual and somatic functions in the parietal associative area 7 of the monkey. *Brain research*.
- Iachini, T., Coello, Y., Frassinetti, F., & Ruggiero, G. (2014). Body space in social interactions: a comparison of reaching and comfort distance in immersive virtual reality. *PloS one*, 9(11), e111511.
- Iachini, T., Coello, Y., Frassinetti, F., Senese, V. P., Galante, F., and Ruggiero, G. (2016). Peripersonal and interpersonal space in virtual and real environments: effects of gender and age. *J. Environ. Psychol.* 45, 154–164.
- Iachini, T., Pagliaro, S., & Ruggiero, G. (2015). Near or far? It depends on my impression: Moral information and spatial behavior in virtual interactions. *Acta psychologica*, 161, 131-136.
- Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., & Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS biology*, 3(3), e79.
- Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *science*, 286(5449), 2526-2528.
- Iriki, A., Tanaka, M., & Iwamura, Y. (1996). Coding of modified body schema during tool use by macaque postcentral neurones. *Neuroreport*, 7(14), 2325-2330.

- Ishida, H., Nakajima, K., Inase, M., and Murata, A. (2010). Shared mapping of own and others' bodies in visuo-tactile bimodal area of monkey parietal cortex. *J. Cogn. Neurosci.* 22, 83–96.
- Ito, T. A., & Bartholow, B. D. (2009). The neural correlates of race. *Trends in cognitive sciences*, 13(12), 524-531.
- Ito, T. A., & Urland, G. R. (2005). The influence of processing objectives on the perception of faces: An ERP study of race and gender perception. *Cognitive, Affective, & Behavioral Neuroscience*, 5(1), 21-36.
- Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological science*, 16(5), 367-371.
- Jacobs, S., Brozzoli, C., Hadj-Bouziane, F., Meunier, M., Farne, A., (2011). Studying multisensory processing and its role in the representation of space through pathological and physiological crossmodal extinction. *Front. Psychol.* 2, 89.
- Jagini, K. K. (2021). Temporal Binding in multisensory and Motor-Sensory contexts: toward a unified model. *Frontiers in Human Neuroscience*, 15. <https://doi.org/10.3389/fnhum.2021.629437>
- Jeannerod, M., and Biguer, B. (1989). Egocentric reference and represented space. *Rev. Neurol.* 145, 635–639.
- Jiang, L. C., Bazarova, N. N., & Hancock, J. T. (2011). The Disclosure–Intimacy Link in Computer-Mediated Communication: An Attributional Extension of the Hyperpersonal Model. *Human Communication Research*, 37(1), 58–77. <https://doi.org/10.1111/j.1468-2958.2010.01393.x>
- Johnson, M., and Demiris, Y. (2005). Perceptual perspective taking and action recognition. *Int. J. Adv. Robot. Syst.* 2, 301–309.
- Jolliffe, I. T. (2010). *Principal component analysis*. Springer.
- Jung, J., Bungert, A., Bowtell, R., & Jackson, S. R. (2016). Vertex stimulation as a control site for transcranial magnetic stimulation: a concurrent TMS/fMRI study. *Brain stimulation*, 9(1), 58-64.
- Jung, T., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2000). Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clinical Neurophysiology*, 111(10), 1745–1758. [https://doi.org/10.1016/s1388-2457\(00\)00386-2](https://doi.org/10.1016/s1388-2457(00)00386-2)
- Kalckert, A., & Ehrsson, H. H. (2012). Moving a rubber hand that feels like your own: a dissociation of ownership and agency. *Frontiers in human neuroscience*, 6, 40.
- Kammers, M. P., de Vignemont, F., Verhagen, L., & Dijkerman, H. C. (2009). The rubber hand illusion in action. *Neuropsychologia*, 47(1), 204-211.
- Kassambara, A. (2020). rstatix: Pipe-Friendly Framework for Basic Statistical Tests. R package version 0.7.0.
- Keller, P. E., Novembre, G., and Hove, M. J. (2014). Rhythm in joint action: psychological and neurophysiological mechanisms for real-time interpersonal coordination. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369:20130394.
- Kelso, J. A. S., Dumas, G., and Tognoli, E. (2013). Outline of a general theory of behavior and brain coordination. *Neural Netw.* 37, 120–131.
- Kemp, A. H., & Quintana, D. S. (2013). The relationship between mental and physical health: insights from the study of heart rate variability. *International journal of Psychophysiology*, 89(3), 288-296.
- Keogh, R., & Pearson, J. (2011). Mental imagery and visual working memory. *PloS one*, 6(12), e29221.
- Keogh, R., & Pearson, J. (2014). The sensory strength of voluntary visual imagery predicts visual working memory capacity. *Journal of Vision*, 14(12), 7–7.
- Kessler, K., Biermann-Ruben, K., Jonas, M., Roman Siebner, H., Bäumer, T., Münchau, A., et al. (2006). Investigating the human mirror neuron system by means of cortical synchronization during the imitation of biological movements. *Neuroimage* 33, 227–238.
- Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cognitive processing*, 8, 159-166.
- Kilner, J. M., Vargas, C., Duval, S., Blakemore, S. J., & Sirigu, A. (2004). Motor activation prior to observation of a predicted movement. *Nature neuroscience*, 7(12), 1299-1301.
- Kilteni, K., & Ehrsson, H. H. (2022). Predictive attenuation of touch and tactile gating are distinct perceptual phenomena. *iScience*, 25(4), 104077. <https://doi.org/10.1016/j.isci.2022.104077>
- King, J., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: the temporal generalization method. *Trends in Cognitive Sciences*, 18(4), 203–210. <https://doi.org/10.1016/j.tics.2014.01.002>
- Knoblich, G., & Jordan, J. S. (2003). Action coordination in groups and individuals: learning anticipatory control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(5), 1006.
- Kokal, I., Gazzola, V., & Keysers, C. (2009). Acting together in and beyond the mirror neuron system. *Neuroimage*, 47(4), 2046-2056.
- Konvalinka, I., Roepstorff, A. (2012). The two-brain approach: how can mutually interacting brains teach us something about social interaction? *Frontiers in Human Neuroscience*, 6, 215.
- Konvalinka, I., Vuust, P., Roepstorff, A., Frith, C.D. (2010). Follow you, follow me: continuous mutual prediction and adaptation in joint tapping. *The Quarterly Journal of Experimental Psychology*, 63(11), 2220–30.
- Körding, K. P., Beierholm, U., Ji, W., MA, Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal inference in multisensory perception. *PLoS ONE*, 2(9), e943. <https://doi.org/10.1371/journal.pone.0000943>

- Kourtis, D., Sebanz, N., & Knoblich, G. (2013). Predictive representation of other people's actions in joint action planning: An EEG study. *Social neuroscience*, 8(1), 31-42.
- Kubota, J. T., Banaji, M. R. & Phelps, E. A. (2012). The neuroscience of race. *Nat. Neurosci.* 15, 940–8.
- Kubota, J. T., Banaji, M. R., & Phelps, E. A. (2012). The neuroscience of race. *Nature neuroscience*, 15(7), 940-948.
- Kuhn, M., & Wickham, H. (2018). tidymodels: Easily install and load the ‘tidymodels’ packages. Retrieved from <https://CRAN.R-project.org/package=tidymodels>
- Làdavas, E. (2002). Functional and dynamic properties of visual peripersonal space. *Trends in cognitive sciences*, 6(1), 17-22.
- Làdavas, E., & Farnè, A. (2004). Visuo-tactile representation of near-the-body space. *Journal of Physiology-Paris*, 98(1-3), 161-170.
- Làdavas, E., & Serino, A. (2010). How the Sense of Body Influences the Sense of Touch. *The cognitive neuroscience of mind: A tribute to Michael S. Gazzaniga*, 155.
- Lahnakoski, J. M., Forbes, P. A. G., McCall, C., & Schilbach, L. (2020). Unobtrusive tracking of interpersonal orienting and distance predicts the subjective quality of social interactions. *Royal Society Open Science*, 7(8), 191815.
- Larsson, D. E., Esposito, G., Critchley, H. D., Dienes, Z., & Garfinkel, S. N. (2021). Sensitivity to changes in rate of heartbeats as a measure of interoceptive ability. *Journal of Neurophysiology*, 126(5), 1799-1813.
- Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software*, 25(1), 1–18.
- Legrand, D., Brozzoli, C., Rossetti, Y., & Farnè, A. (2007). Close to me: Multisensory space representations for action and pre-reflexive consciousness of oneself-in-the-world. *Consciousness and cognition*, 16(3), 687-699.
- Leibman, M. (1970). The effects of sex and race norms on personal space. *Environ. Behav.* 2, 208–246.
- Leinonen, L. E. A. (1980). Functional properties of neurones in the posterior part of area 7 in awake monkey. *Acta Physiologica Scandinavica*, 108(3), 301-308.
- Leinonen, L., & Nyman, G. (1979). II. Functional properties of cells in anterolateral part of area 7 associative face area of awake monkeys. *Experimental Brain Research*, 34, 321-333.
- Lenth, R. V. (2020). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.4.8.
- Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2019). Package ‘emmeans’.
- Lewis, M., Colunga, E., & Lupyan, G. (2021). Superordinate word knowledge predicts longitudinal vocabulary growth. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 43, No. 43).
- Liberman, N., Trope, Y., & Stephan, E. (2007). Psychological distance. In *Social psychology: Handbook of basic principles*, 2nd ed (pp. 353–381). The Guilford Press
- Liew, S.-L., Han, S., & Aziz-Zadeh, L. (2011). Familiarity modulates mirror neuron and mentalizing regions during intention understanding. *Human Brain Mapping*, 32(11), 1986–1997.
- Link, B. G., & Phelan, J. C. (2001). Conceptualizing stigma. *Annual review of Sociology*, 27(1), 363-385.
- Lischke, A., Pahnke, R., Mau-Moeller, A., & Weippert, M. (2021). Heart rate variability modulates interoceptive accuracy. *Frontiers in Neuroscience*, 14, 612445.
- Lisi, M. P., Scattolin, M., Fusaro, M., & Aglioti, S. M. (2021). A Bayesian approach to reveal the keyrole of mask wearing in modulating projected interpersonal distance during the first COVID-19outbreak. *PLOS ONE*, 16(8), e0255598.
- Liu, D., Baumeister, R. F., Yang, C., & Hu, B. (2019). Retracted: Digital Communication Media Useand Psychological Well-Being: A Meta-Analysis. *Journal of Computer-Mediated Communication*, 24(5), 259–273.
- Llobera, J., Charbonnier, C., Chagué, S., Preissmann, D., Antonietti, J.-P., Ansermet, F., et al. (2016). The subjective sensation of synchrony: an experimental study. *PLoS One* 11:e0147008.
- Lloyd, D. M. (2007). Spatial limits on referred touch to an alien limb may reflect boundaries of visuo-tactile peripersonal space surrounding the hand. *Brain and cognition*, 64(1), 104-109.
- Lockard, J. S., Mccvittie, R. I., and Isaac, L. M. (1977). Functional significance of the affiliative smile. *Bull. Psychon. Soc.* 9, 367–370.
- Longo, M. R., Schüür, F., Kammers, M. P., Tsakiris, M., & Haggard, P. (2008). What is embodiment? A psychometric approach. *Cognition*, 107(3), 978-998.
- Losin, E. A. R., Iacoboni, M., Martin, A., Cross, K. A., & Dapretto, M. (2012). Race modulates neural activity during imitation. *NeuroImage*, 59(4), 3594–3603.
- Lübbert, A., Göschl, F., Krause, H., Schneider, T. R., Maye, A., & Engel, A. K. (2021). Socializing sensorimotor contingencies. *Frontiers in Human Neuroscience*, 15, 624610.
- Lupyan, G. (2012). Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in psychology*, 3, 54.
- Lupyan, G., & Winter, B. (2018). Language is more abstract than you think, or, why aren’t languages more iconic? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1752), 20170137.

- Lynott, D., & Connell, L. (2013). Modality exclusivity norms for 400 nouns: The relationship between perceptual experience and surface word form. *Behavior research methods*, pp. 45, 516–526.
- Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior research methods*, 47, 1122-1135.
- Ma, N., Baetens, K., Vandekerckhove, M., Van der Cruyssen, L., & Van Overwalle, F. (2013). Dissociation of a trait and a valence representation in the mPFC. *Social Cognitive and Affective Neuroscience*, 9(10), 1506–1514.
- Macaluso, E., & Maravita, A. (2010). The representation of space near the body through touch and vision. *Neuropsychologia*, 48(3), 782-795.
- Mackey, S., & Petrides, M. (2014). Architecture and morphology of the human ventromedial prefrontal cortex. *European Journal of Neuroscience*, 40(5), 2777-2796.
- Mackie, D. M., & Smith, E. R. (2002). Intergroup emotions and the social self: Prejudice reconceptualized as differentiated reactions to out-groups. In J. P. Forgas & K. D. Williams (Eds.), *The social self: Cognitive, interpersonal and intergroup perspectives* (pp. 309–326). New York, NY: Psychology Press.
- Magosso, E., Serino, A., Di Pellegrino, G., & Ursino, M. (2010b). Crossmodal links between vision and touch in spatial attention: a computational modelling study. *Computational Intelligence and Neuroscience*, 2010(1), 304941.
- Magosso, E., Zavaglia, M., Serino, A., Di Pellegrino, G., & Ursino, M. (2010a). Visuo-tactile representation of peripersonal space: a neural network study. *Neural computation*, 22(1), 190-243.
- Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of physiology-Paris*, 102(1-3), 59-70.
- Maister, L., Sebanz, N., Knoblich, G., & Tsakiris, M. (2013). Experiencing ownership over a dark-skinned body reduces implicit racial bias. *Cognition*, 128(2), 170-178.
- Maister, L., Slater, M., Sanchez-Vives, M. V., & Tsakiris, M. (2015). Changing bodies changes minds: owning another body affects social cognition. *Trends in cognitive sciences*, 19(1), 6-12.
- Makin, T. R., Holmes, N. P., & Ehrsson, H. H. (2008). On the other hand: dummy hands and peripersonal space. *Behavioural brain research*, 191(1), 1-10.
- Makin, T. R., Holmes, N. P., & Zohary, E. (2007). Is that near my hand? Multisensory representation of peripersonal space in human intraparietal sulcus. *Journal of Neuroscience*, 27(4), 731-740.
- Makin, T. R., Holmes, N. P., Brozzoli, C., & Farnè, A. (2012). Keeping the world at hand: rapid visuomotor processing for hand–object interactions. *Experimental brain research*, 219, 421-428.
- Maravita, A. (2006). From “body in the brain” to “body in space”: sensory and intentional components of body representation. *Human body perception from the inside out*. Oxford University Press, New York, 65-88.
- Maravita, A., & Iriki, A. (2004). Tools for the body (schema). *Trends in cognitive sciences*, 8(2), 79-86.
- Maravita, A., Husain, M., & Clarke, K. (2003b). Reaching with a tool extends visual-tactile interactions into far space: Evidence from cross-modal extinction. *Neuropsychologia*, 39(6), 580-585.
- Maravita, A., Spence, C., & Driver, J. (2003a). Multisensory integration and the body schema: close to hand and within reach. *Current biology*, 13(13), R531-R539.
- Maravita, A., Spence, C., Kennett, S., & Driver, J. (2002). Tool-use changes multimodal spatial interactions between vision and touch in normal humans. *Cognition*, 83(2), B25-B34.
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164(1), 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>
- Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, 64(1), 17–24.
- Mattingley, J.B., Driver, J., Beschin, N., Robertson, I.H., 1997. Attentional competition between modalities: extinction between touch and vision after right hemisphere damage. *Neuropsychologia* 35 (6), 867–880.
- Maye, A., & Engel, A. K. (2011). A discrete computational model of sensorimotor contingencies for object perception and control of behavior. In *2011 IEEE International Conference on Robotics and Automation* (pp. 3810-3815). IEEE.
- Maye, A., & Engel, A. K. (2012). Using sensorimotor contingencies for prediction and action planning. In *From Animals to Animats 12: 12th International Conference on Simulation of Adaptive Behavior, SAB 2012, Odense, Denmark, August 27-30, 2012. Proceedings 12* (pp. 106-116). Springer Berlin Heidelberg.
- Mazzuca, C., & Santarelli, M. (2022). Making it abstract, making it contestable: Politicization at the intersection of political and cognitive science. *Review of Philosophy and Psychology*, 1–22.
- Mazzuca, C., Falcinelli, I., Michalland, A. H., Tummolini, L., & Borghi, A. M. (2022). Bodily, emotional, and public sphere at the time of COVID-19. An investigation on concrete and abstract concepts. *Psychological Research*, 86(7), 2266–2277.
- Mazzuca, C., Fini, C., Michalland, A. H., Falcinelli, I., Da Rold, F., Tummolini, L., & Borghi, A. M. (2021). From affordances to abstract words: The flexibility of sensorimotor grounding. *Brain Sciences*, 11(10), 1304.
- Mazzuca, C., Lugli, L., Benassi, M., Nicoletti, R., & Borghi, A. M. (2018). Abstract, emotional and concrete concepts and the activation of mouth-hand effectors. *PeerJ*, 6, e5987.

- Mazzuca, C., Majid, A., Lugli, L., Nicoletti, R., & Borghi, A. M. (2020). Gender is a multifaceted concept: Evidence that specific life experiences differentially shape the concept of gender. *Language and Cognition*, 12(4), 649–678.
- McKenna, K. Y. A., Green, A. S., & Gleason, M. E. J. (2002). Relationship formation on the Internet: What's the big attraction? *Journal of Social Issues*, 58(1), 9–31. <https://doi.org/10.1111/1540-4560.00246>
- McKinney, W. (2010, June). Data structures for statistical computing in Python. In *SciPy* (Vol. 445, No. 1, pp. 51-56).
- Meltzoff, A.N. and Decety, J. (2003) What imitation tells us about social cognition: a rapprochement between developmental psychology and cognitive neuroscience. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 491–500
- Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Networks*, 9, 1265–1279.
- Miller, M. R., Jun, H., Herrera, F., Villa, J. Y., Welch, G., & Bailenson, J. N. (2019). Social interaction in augmented reality. *PLOS ONE*, 14(5), e0216290. <https://doi.org/10.1371/journal.pone.0216290>
- Milner, A. D., & Goodale, M. A. (2008). Two visual systems re-viewed. *Neuropsychologia* 46, 774–785.
- Milner, D., & Goodale, M. (1995). *The visual brain in action*. Oxford University Press, Oxford
- Mitkidis, P., McGraw, J.J., Roepstorff, A., Wallot, S. (2015). Building trust: heart rate synchrony and arousal during joint action increased by public goods game. *Physiology & Behavior*, 149, 101–6.
- Mohammad, S. (2018). Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 English words. In *Proceedings of the 56th annual meeting of the association for computational linguistics (volume 1: long papers)* (pp. 174–184).
- Molnar-Szakacs, I., Wu, A. D., Robles, F. J., & Iacoboni, M. (2007). Do you see what I mean? Corticospinal excitability during observation of culture-specific gestures. *PloS one*, 2(7), e626.
- Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2014). The adaptation of the affective norms for English words (ANEW) for Italian. *Behavior Research Methods*, 46(3), 887–903.
- Moreau, Q., Candidi, M., Era, V., Tieri, G., & Aglioti, S. M. (2020). Midline frontal and occipito-temporal activity during error monitoring in dyadic motor interactions. *Cortex*, 127, 131-149.
- Morrison, K. E., DeBrabander, K. M., Jones, D. R., Faso, D. J., Ackerman, R. A., & Sasson, N. J. (2020). Outcomes of real-world social interaction for autistic adults paired with autistic compared to typically developing partners. *Autism: The International Journal of Research and Practice*, 24(5), 1067–1080.
- Muller, V., Lindenberger, U. (2011). Cardiac and respiratory patterns synchronize between persons during choir singing. *PLoS ONE*, 6(9), e24893.
- Muraki, E. J., Sidhu, D. M., & Pexman, P. M. (2022). Heterogenous abstract concepts: Is ‘ponder’ different from ‘dissolve’? *Psychological Research*, 86(8), 2478–2494. <https://doi.org/10.1007/s00426-020-01398-x>
- Muraki, E. J., Speed, L. J., & Pexman, P. M. (2023). Insights into embodied cognition and mental imagery from aphantasia. *Nature Reviews Psychology*, 2(10), 591-605.
- Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: affective priming with optimal and suboptimal stimulus exposures. *Journal of personality and social psychology*, 64(5), 723.
- Muscattell, K. A., Morelli, S. A., Falk, E. B., Way, B. M., Pfeifer, J. H., Galinsky, A. D., ... & Eisenberger, N. I. (2012). Social status modulates neural activity in the mentalizing network. *Neuroimage*, 60(3), 1771-1777.
- Myers, M. W., & Hodges, S. D. (2012). The structure of self–other overlap and its relationship to perspective taking. *Personal Relationships*, 19(4), 663-679.
- Naselaris, T., Olman, C. A., Stansbury, D. E., Ugurbil, K., & Gallant, J. L. (2015). A voxel-wise encoding model for early visual areas decodes mental images of remembered scenes. *Neuroimage*, 105, 215-228.
- Neal, A., and Kilner, J. M. (2010). What is simulated in the action observation network when we observe actions? *Eur. J. Neurosci.* 32, 1765–1770.
- Newman-Norlund, R. D., van Schie, H. T., van Zuijlen, A. M., & Bekkering, H. (2007). The mirror neuron system is more active during complementary compared with imitative action. *Nature neuroscience*, 10(7), 817-818.
- Nguyen, M., Bin, Y. S., & Campbell, A. (2012). Comparing online and offline self-disclosure: A systematic review. *Cyberpsychology, Behavior and Social Networking*, 15(2), 103–111.
- Noë, A. (2004). *Action in Perception*. Cambridge, MA: MIT Press.
- Noel, J. P., Grivaz, P., Marmaroli, P., Lissek, H., Blanke, O., & Serino, A. (2015). Full body action remapping of peripersonal space: the case of walking. *Neuropsychologia*, 70, 375-384.
- Noppeney, U. (2021). Perceptual inference, learning, and attention in a multisensory world. *Annual Review of Neuroscience*, 44(1), 449–473. <https://doi.org/10.1146/annurev-neuro-100120-085519>
- Norman, E., Tjomsland, H. E., & Huegel, D. (2016). The Distance between Us: Using Construal Level Theory to Understand Interpersonal Distance in a Digital Age. *Frontiers in Digital Humanities*, 3.
- Novembre, G., Knoblich, G., Dunne, L., Keller, P.E. (2017). Interpersonal synchrony enhanced through 20 Hz phase-coupled dual brain stimulation. *Social Cognitive and Affective Neuroscience*, 12(4), 662–70.
- Novembre, G., Ticini, L.F., Schütz-Bosbach, S., Keller, P.E. (2014). Motor simulation and the coordination of self and other in real-time joint action. *Social Cognitive and Affective Neuroscience*, 9(8), 1062–8.

- O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24(5), 939–973. <https://doi.org/10.1017/s0140525x01000115>
- Ocampo, B., & Kritikos, A. (2010). Placing actions in context: motor facilitation following observation of identical and non-identical manual acts. *Experimental brain research*, 201, 743-751.
- Oh, H., Curley, S., & Subramani, M. (2008). The Death of Distance?: The Influence of ComputerMediated Communication on Perceptions of Distance. ICIS 2008 Proceedings. <https://aisel.aisnet.org/icis2008/149>
- Okamoto, M., Dan, H., Sakamoto, K., Takeo, K., Shimizu, K., Kohno, S., ... & Dan, I. (2004). Three-dimensional probabilistic anatomical cranio-cerebral correlation via the international 10–20 system oriented for transcranial functional brain mapping. *Neuroimage*, 21(1), 99-111.
- Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, 2011, 1–9. <https://doi.org/10.1155/2011/156869>
- Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, 105(32), 11087-11092.
- O'regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and brain sciences*, 24(5), 939-973.
- Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review*, 76(3),241–263.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.
- Paivio, A. (1990). *Mental representations: A dual coding approach*. Oxford University Press.
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 45(3), 255.
- Paivio, A., Clark, J. M., & Lambert, W. E. (1988). Bilingual dual-coding theory and semantic repetition effects on recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(1), 163.
- Parks, M. R., & Floyd, K. (1996). Making Friends in Cyberspace. *Journal of Computer-Mediated Communication*, 1(4), 0–0.
- Pascual-Leone, A., Walsh, V., & Rothwell, J. (2000). Transcranial magnetic stimulation in cognitive neuroscience—virtual lesion, chronometry, and functional connectivity. *Current opinion in neurobiology*, 10(2), 232-237.
- Patané, I., Brozzoli, C., Koun, E., Frassinetti, F., & Farnè, A. (2021). Me, you, and our object: Peripersonal space recruitment during executed and observed actions depends on object ownership. *Journal of experimental psychology: general*, 150(7), 1410.
- Patané, I., Cardinali, L., Salemme, R., Pavani, F., Farnè, A., & Brozzoli, C. (2019). Action planning modulates peripersonal space. *Journal of Cognitive Neuroscience*, 31(8), 1141-1154.
- Patané, I., Farnè, A., & Frassinetti, F. (2017). Cooperative tool-use reveals peripersonal and interpersonal spaces are dissociable. *Cognition*, 166, 13-22.
- Patané, I., Iachini, T., Farnè, A., & Frassinetti, F. (2016). Disentangling action from social space: tool-use differently shapes the space around us. *PLoS one*, 11(5), e0154247.
- Pavan, G., Dalmaso, M., Galfano, G. & Castelli, L. (2011). Racial group membership is associated to gaze-mediated orienting in Italy. *PLoS One* 6, e25608
- Pavani, F., & Castiello, U. (2004). Binding personal and extrapersonal space through body shadows. *Nature neuroscience*, 7(1), 14-16.
- Pavani, F., Spence, C., & Driver, J. (2000). Visual capture of touch: Out-of-the-body experiences with rubber gloves. *Psychological science*, 11(5), 353-359.
- Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: affect misattribution as implicit measurement. *Journal of personality and social psychology*, 89(3), 277.
- Payne, K., & Lundberg, K. (2014). The affect misattribution procedure: Ten years of evidence on reliability, validity, and mechanisms. *Social and Personality Psychology Compass*, 8(12), 672-686.
- Pearson, J., Naselaris, T., Holmes, E. A., & Kosslyn, S. M. (2015). Mental imagery: functional mechanisms and clinical applications. *Trends in cognitive sciences*, 19(10), 590–602.
- Peck, T. C., Seinfeld, S., Aglioti, S. M., & Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and cognition*, 22(3), 779-787.
- Pedersen T (2022). `_patchwork: The Composer of Plots_`. R package version 1.1.2, <https://CRAN.R-project.org/package=patchwork>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *The Journal of machine Learning research*, 12, 2825-2830.
- Pellencin, E., Paladino, M. P., Herbelin, B., and Serino, A. (2017). Social perception of others shapes one's own multisensory peripersonal space. *Cortex* 104,163–179.
- Petitto, L. A., & Marentette, P. F. (1991). Babbling in the manual mode: Evidence for the ontogeny of language. *Science*, 251(5000), 1493-1496.

- Pettegrew, L. S., & Day, C. (2015). Smart Phones and Mediated Relationships: The Changing Face of Relational Communication. *Review of Communication*, 15(2), 122–139. <https://doi.org/10.1080/15358593.2015.1044018>
- Pexman, P. M., Diveica, V., & Binney, R. J. (2023). Social semantics: the organization and grounding of abstract concepts. *Philosophical Transactions of the Royal Society B*, 378(1870), 20210363
- Pezzulo, G., Iodice, P., Donnarumma, F., Dindo, H., & Knoblich, G. (2017). Avoiding accidents at the champagne reception: A study of joint lifting and balancing. *Psychological science*, 28(3), 338-345.
- Pezzulo, G., Iodice, P., Ferraina, S., & Kessler, K. (2013). Shared action spaces: a basis function framework for social re-calibration of sensorimotor representations supporting joint action. *Frontiers in human neuroscience*, 7, 800.
- Pfeifer, R., and Bongard, J. C. (2006). *How the Body Shapes the Way We Think: A New View of Intelligence*. Cambridge, MA: MIT Press.
- Politzer-Ahles, S., & Piccinini, P. (2018). On visualizing phonetic data from repeated measures experiments with multiple random effects. *Journal of Phonetics*, 70, 56–69. <https://doi.org/10.1016/j.wocn.2018.05.002>
- Pollick, A. S., & De Waal, F. B. (2007). Ape gestures and language evolution. *Proceedings of the National Academy of Sciences*, 104(19), 8184-8189.
- Ponari, M., Norbury, C. F., & Vigliocco, G. (2018). Acquisition of abstract concepts is influenced by emotional valence. *Developmental Science*, 21(2). <https://doi.org/10.1111/desc.12549>
- Pouget, A., Deneve, S., and Duhamel, J.-R. (2002). A computational perspective on the neural basis of multisensory spatial representations. *Nat. Rev. Neurosci.* 3, 741–747.
- Powers, K. E., Wagner, D. D., Norris, C. J., & Heatherton, T. F. (2013). Socially excluded individuals fail to recruit medial prefrontal cortex for negative social scenes. *Social Cognitive and Affective Neuroscience*, 8(2), 151–157.
- Prinz, J. J. (2012). *Beyond human nature. How culture and experience shape our lives*. Penguin.
- Prinz, W. (1997). Perception and action planning. *Eur. J. Cogn. Psychol.* 9, 129–154.
- R Core Team (2019) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological bulletin*, 114(3), 510.
- Reggin, L. D., Muraki, E. J., & Pexman, P. M. (2021). Development of abstract word knowledge. *Frontiers in Psychology*, 12, 2115
- Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive science*, 29(6), 1045-1060.
- Richardson, M.J., Marsh, K.L., Isenhower, R.W., Goodman, J.R.L., Schmidt, R.C. (2007). Rocking together: dynamics of intentional and unintentional interpersonal coordination. *Human Movement Science*, 26(6), 867–91.
- Riemer, M., Kleinböhl, D., Hölzl, R., & Trojan, J. (2013). Action and perception in the rubber hand illusion. *Experimental brain research*, 229, 383-393.
- Rietveld, E., & Kiverstein, J. (2014). A rich landscape of affordances. *Ecological psychology*, 26(4), 325-352.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.*, 27(1), 169-192.
- Rizzolatti, G., & Fogassi, L. (2014). The mirror mechanism: Recent findings and perspectives. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 369, 20130420.
- Rizzolatti, G., & Gentilucci, M. (1988). Motor and visual-motor functions of the premotor cortex. *Neurobiology of neocortex*, 42, 269-284.
- Rizzolatti, G., & Semigallia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature reviews neuroscience*, 11(4), 264-274.
- Rizzolatti, G., & Sinigaglia, C. (2016). The mirror mechanism: A basic principle of brain function. *Nature Reviews Neuroscience*, 17, 757–765. <https://doi.org/10.1038/nrn.2016.135>
- Rizzolatti, G., and Luppino, G. (2001). The cortical motor system. *Neuron* 31, 889–901.
- Rizzolatti, G., Fadiga, L., Fogassi, L., and Gallese, V. (1997). The space around us. *Science* 277, 190–191.
- Rizzolatti, G., Gentilucci, M., Fogassi, L., Luppino, G., Matelli, M., & Ponzoni-Maggi, S. (1987). Neurons related to goal-directed motor acts in inferior area 6 of the macaque monkey. *Experimental Brain Research*, 67(1), 220-224.
- Rizzolatti, G., Luppino, G., & Matelli, M. (1998). The organization of the cortical motor system: new concepts. *Electroencephalography and clinical neurophysiology*, 106(4), 283-296.
- Rizzolatti, G., Riggio, L., & Sheliga, B. M. (1994). Space and selective attention. *Attention and performance XV*, 15, 231-265.
- Rizzolatti, G., Scandolara, C., Matelli, M. and Gentilucci, M. (1981a). Afferent properties of periarculate neurons in macaque monkeys. I. Somatosensory responses. *Behavioral Brain Research*, 2, 125-46.
- Rizzolatti, G., Scandolara, C., Matelli, M. and Gentilucci, M. (1981b). Afferent properties of periarculate neurons in macaque monkeys. II. Visual responses. *Behavioral Brain Research*, 2, 147-163.
- Rock, I., & Victor, J. (1964). Vision and touch: An experimentally created conflict between the two senses. *Science*, 143(3606), 594-596.
- Rohde, M., Di Luca, M., & Ernst, M. O. (2011). The rubber hand illusion: feeling of ownership and proprioceptive drift do not go hand in hand. *PloS one*, 6(6), e21659.

- Rohe, T., Ehrlis, A. C., & Noppeney, U. (2019). The neural dynamics of hierarchical Bayesian causal inference in multisensory perception. *Nature communications*, 10(1), 1907.
- Ronga, I., Galigani, M., Bruno, V., Noel, J., Gazzin, A., Perathoner, C., Serino, A., & Garbarini, F. (2021). Spatial tuning of electrophysiological responses to multisensory stimuli reveals a primitive coding of the body boundaries in newborns. *Proceedings of the National Academy of Sciences of the United States of America*, 118(12).
- Rosenthal, C.R., Roche-Kelly, E.E., Husain, M., Kennard, C. (2009). Response-dependent contributions of human primary motor cortex and angular gyrus to manual and perceptual sequence learning. *Journal of Neuroscience*, 29(48), 15115–25.
- Rossini, P. M., Burke, D., Chen, R., Cohen, L. G., Daskalakis, Z., Di Iorio, R., ... & Ziemann, U. (2015). Non-invasive electrical and magnetic stimulation of the brain, spinal cord, roots and peripheral nerves: Basic principles and procedures for routine clinical and research application. An updated report from an IFCN Committee. *Clinical neurophysiology*, 126(6), 1071-1107.
- Ruggiero, G., Frassinetti, F., Coello, Y., Rapuano, M., Di Cola, A. S., and Iachini, T. (2017). The effect of facial expressions on peripersonal and interpersonal spaces. *Psychol. Res.* 81, 1232–1240.
- Rumiati, R. I., & Foroni, F. (2016). We are what we eat: How food is represented in our mind/brain. *Psychonomic Bulletin & Review*, 23(4), 1043–1054.
- Ryan, T., Allen, K. A., Gray, D. L., & McInerney, D. M. (2017). How Social Are Social Media? A Review of Online Social Behaviour and Connectedness. *Journal of Relationships Research*, 8,e8. <https://doi.org/10.1017/jrr.2017.13>
- Sacheli, L. M., Aglioti, S. M., & Candidi, M. (2015a). Social cues to joint actions: The role of shared goals. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01034>
- Sacheli, L. M., Arcangeli, E., & Paulesu, E. (2018). Evidence for a dyadic motor plan in joint action. *Scientific reports*, 8(1), 5027.
- Sacheli, L. M., Candidi, M., Era, V., & Aglioti, S. M. (2015b). Causative role of left aIPS in coding shared goals during human–avatar complementary joint actions. *Nature Communications*, 6(1). <https://doi.org/10.1038/ncomms8544>
- Sacheli, L. M., Candidi, M., Pavone, E. F., Tidoni, E., & Aglioti, S. M. (2012). And yet they act together: Interpersonal perception modulates visuo-motor interference and mutual adjustments during a joint-grasping task. *PLoS ONE*, 7(11). <https://doi.org/10.1371/journal.pone.0050223>
- Sacheli, L. M., Tidoni, E., Pavone, E. F., Aglioti, S. M., & Candidi, M. (2013). Kinematics fingerprints of leader and follower role-taking during cooperative joint actions. *Experimental Brain Research*, 226(4), 473–486.
- Schilbach, L. (2015). Eye to eye, face to face and brain to brain: novel approaches to study the behavioral dynamics and neural mechanisms of social interactions. *Current Opinion in Behavioral Sciences*, 3, 130-135.
- Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences*, 4, 393–414.
- Schilbach, L., Wilms, M., Eickhoff, S. B., Romanzetti, S., Tepest, R., Bente, G., ... & Vogeley, K. (2010). Minds made for sharing: initiating joint attention recruits reward-related neurocircuitry. *Journal of cognitive neuroscience*, 22(12), 2702-2715.
- Schmitz, L., & Reader, A. T. (2023). Smaller preferred interpersonal distance for joint versus parallel action. *PLOS ONE*, 18(5), e0285202. <https://doi.org/10.1371/journal.pone.0285202>
- Schwanenflugel, P. J. (1991). Contextual constraint and lexical processing. *Advances in psychology* (Vol. 77, pp. 23-45). North-Holland.
- Schwanenflugel, P. J., Akin, C., & Luh, W. M. (1992). Context availability and the recall of abstract and concrete words. *Memory & Cognition*, 20(1), 96-104.
- Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language*, 27(5), 499–520.
- Sciortino, P., & Kayser, C. (2022). Multivariate Analysis of Evoked Responses during the Rubber Hand Illusion Suggests a Temporal Parcellation into Manipulation and Illusion-Specific Correlates. *ENeuro*, 9(1), ENEURO.0355-21.2021.
- Sebanz, N. and Frith, C. (2004) Beyond simulation? Neural mechanisms for predicting the actions of others. *Nat. Neurosci.* 7, 5–6
- Sebanz, N., and Knoblich, G. (2009). Prediction in joint action: what, when, and where. *Top. Cogn. Sci.* 1, 353–367. doi: 10.1111/j.1756-8765.2009.01024.x
- Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in cognitive sciences*, 10(2), 70-76.
- Sebanz, N., Knoblich, G., & Prinz, W. (2005). How two share a task: corepresenting stimulus-response mappings. *Journal of Experimental Psychology: Human Perception and Performance*, 31(6), 1234.
- Sellaro, R., Derks, B., Nitsche, M. A., Hommel, B., van den Wildenberg, W. P., van Dam, K., & Colzato, L. S. (2015). Reducing prejudice through brain stimulation. *Brain stimulation*, 8(5), 891-897.
- Senna, I., Cardinali, L., Farnè, A., & Brozzoli, C. (2019). Aim and plausibility of action chains remap peripersonal space. *Frontiers in Psychology*, 10, 1681.

- Sereno, M.I. & Huang, R.S. (2006). A human parietal face area contains aligned head centered visual and tactile maps. *Nature Neuroscience*, 9: 1337-43.
- Serino, A. (2019). Peripersonal space (PPS) as a multisensory interface between the individual and the environment, defining the space of the self. *Neuroscience & Biobehavioral Reviews*, 99, 138-159.
- Serino, A., Bassolino, M., Farne, A., Ladavas, E., (2007). Extended multisensory space in blind cane users. *Psychol. Sci.* 18 (7), 642–648.
- Serino, A., Canzoneri, E., Marzolla, M., di Pellegrino, G., & Magosso, E. (2015). Extending peripersonal space representation without tool-use: evidence from a combined behavioral-computational approach. *Frontiers in Behavioral Neuroscience*, 9.
- Serino, A., Pizzoferrato, F., & Ladavas, E. (2008). Viewing a face (especially one's own face) being touched enhances tactile perception on the face. *Psychological science*, 19(5), 434-438.
- Shadmehr R, Krakauer JW (2008). A computational neuroanatomy for motor control. *Exp Brain Res* 185:359-381.
- Shafaei, F. S., Mirghafourvand, M., & Havizari, S. (2020). The effect of prenatal counseling on breast feeding self-efficacy and frequency of breastfeeding problems in mothers with previous unsuccessful breastfeeding: A randomized controlled clinical trial. *BMC Women's Health*, 20(1),94.
- Shayo, M., & Zussman, A. (2011). Judicial ingroup bias in the shadow of terrorism. *The Quarterly journal of economics*, 126(3), 1447-1484.
- Shea, N. (2018). Metacognition and abstract concepts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1752), 20170133.
- Sheng, F., & Han, S. (2012). Manipulations of cognitive strategies and intergroup relationships reduce the racial bias in empathic neural responses. *NeuroImage*, 61(4), 786-797.
- Shockley, K., Richardson, D. C., & Dale, R. (2009). Conversation and Coordinative Structures. *Topics in Cognitive Science*, 1(2), 305–319. <https://doi.org/10.1111/j.1756-8765.2009.01021.x>
- Siebnner, H., & Rothwell, J. (2003). Transcranial magnetic stimulation: new insights into representational cortical plasticity. *Experimental brain research*, 148, 1-16.
- Siegmán, A. W., and Feldstein, S. (2014). *Nonverbal Behavior and Communication*. New York, NY: *Psychology Press*.
- Simmons, W.K.; Ramjee, V.; Beauchamp, M.S.; McRae, K.; Martin, A.; Barsalou, L.W. A (2007). Common neural substrate for perceiving and knowing about color. *Neuropsychologia*, 45, 2802–2810
- Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2015). Package 'afex'.
- Soliman, T. M., Ferguson, R., Dexheimer, M. S., & Glenberg, A. M. (2015). Consequences of joint action: Entanglement with your partner. *Journal of Experimental Psychology. General*, 144(4), 873–888.
- Spears, R., & Lea, M. (1992). Social influence and the influence of the 'social' in computer-mediated communication. In *Contexts of computer-mediated communication* (pp. 30–65).
- Spence, C., & Driver, J. (Eds.). (2004). *Crossmodal space and crossmodal attention*. Oxford University Press.
- Spence, C., Nicholls, M. E. R., Gillespie, N., & Driver, J. (1998). Cross-modal links in exogenous covert spatial orienting between touch, audition, and vision. *Perception & Psychophysics*, 60(4), 544–557.
- Spence, C., Pavani, F., & Driver, J. (2000). Crossmodal links between vision and touch in covert endogenous spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, 26(4), 1298.
- Spence, C., Pavani, F., & Driver, J. (2004). Spatial constraints on visual-tactile cross-modal distractor congruency effects. *Cognitive, Affective, & Behavioral Neuroscience*, 4(2), 148-169.
- Spence, C., Pavani, F., Maravita, A. and Holmes, N. P. (2008). Multisensory interactions. In *Haptic Rendering: Foundations, Algorithms, and Applications*, eds M.C. Lin, and M.A. Otaduy, 21-52. A K Peters Ltd., Wellesley, MA.
- Spengler, S., von Cramon, D. Y., & Brass, M. (2010). Resisting motor mimicry: control of imitation involves processes central to social cognition in patients with frontal and temporo-parietal lesions. *Social Neuroscience*, 5(4), 401-416.
- Sperry, R. W. (1952). Neurology and the mind-brain problem. *American scientist*, 40(2), 291-312.
- Stanley, D. A., & Adolphs, R. (2013). Toward a neural basis for social behavior. *Neuron*, 80(3), 816-826.
- Steele, J. R., George, M., Cease, M. K., Fabri, T. L., & Schlosser, J. (2018). Not always Black and White: The effect of race and emotional expression on implicit attitudes. *Social Cognition*, 36(5), 534-558.
- Stephan, E., Liberman, N., & Trope, Y. (2010). Politeness and psychological distance: A construal level perspective. *Journal of Personality and Social Psychology*, 98(2), 268–280. <https://doi.org/10.1037/a0016960>
- Stephan, E., Liberman, N., & Trope, Y. (2011). The effects of time perspective and level of construal on social distance. *Journal of Experimental Social Psychology*, 47(2), 397–402.
- Stoet, G. (2010). PsyToolkit - A software package for programming psychological experiments using Linux. *Behavior Research Methods*, 42(4), 1096–1104.
- Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, 44(1), 24–31.

- Stürmer, S., Snyder, M., Kropp, A., & Siem, B. (2006). Empathy-motivated helping: The moderating role of group membership. *Personality and Social Psychology Bulletin*, 32(7), 943-956.
- Tajfel, H. (1982). Experimental studies of intergroup behaviour. In *Cognitive analysis of social behavior: Proceedings of the NATO advanced study Institute on "The cognitive analysis of socio-psychological processes", Aix-en-Provence, France, July 12–31, 1981* (pp. 227-246). Dordrecht: Springer Netherlands.
- Tajfel, H., Billig, M. G., Bundy, R. P., and Flament, C. (1971). Social categorization and intergroup behavior. *Eur. J. Soc. Psychol.* 1, 149–178.
- Talairach, J., & Tournoux, P. (1988). *A Co-Planar Stereotaxic Atlas of the Human Brain* Thieme Medical Publishers: New York. NY.
- Teneggi, C., Canzoneri, E., di Pellegrino, G., & Serino, A. (2013). Social modulation of peripersonal space boundaries. *Current biology*, 23(5), 406-411.
- Teramoto, W. (2018). A behavioral approach to shared mapping of peripersonal space between oneself and others. *Scientific reports*, 8(1), 5432.
- Thomas, R., Press, C., and Haggard, P. (2006). Shared representations in body perception. *Acta Psychol. (Amst.)* 121, 317–330.
- Tidwell, L. C., & Walther, J. B. (2002). Computer-mediated communication effects on disclosure, impressions, and interpersonal evaluations: Getting to know one another a bit at a time. *Human Communication Research*, 28(3), 317–348. <https://doi.org/10.1111/j.1468-2958.2002.tb00811.x>
- Tieri, G., Tidoni, E., Pavone, E. F., & Aglioti, S. M. (2015). Mere observation of body discontinuity affects perceived ownership and vicarious agency over a virtual hand. *Experimental Brain Research*, 233(4), 1247–1259.
- Todorov, A., Mende-Siedlecki, P., & Dotsch, R. (2013). Social judgments from faces. *Current opinion in neurobiology*, 23(3), 373-380.
- Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in cognitive sciences*, 12(12), 455-460.
- Tollefsen, D. (2005). Let's pretend! Children and joint action. *Philosophy of the Social Sciences*, 35(1), 75-97.
- Treder, M. S. (2020). MVPA-Light: A Classification and Regression Toolbox for Multi-Dimensional Data. *Frontiers in Neuroscience*, 14. <https://doi.org/10.3389/fnins.2020.00289>
- Tsakiris, M. (2010). My body in the brain: a neurocognitive model of body-ownership. *Neuropsychologia*, 48(3), 703-712.
- Tsakiris, M., & Haggard, P. (2005). The rubber hand illusion revisited: visuo-tactile integration and self-attribution. *Journal of experimental psychology: Human perception and performance*, 31(1), 80.
- Tsakiris, M., Haggard, P., Franck, N., Mainy, N., & Sirigu, A. (2005). A specific role for efferent information in self-recognition. *Cognition*, 96(3), 215–231.
- Tsakiris, M., Longo, M. R., & Haggard, P. (2010). Having a body versus moving your body: neural signatures of agency and body-ownership. *Neuropsychologia*, 48(9), 2740-2749.
- Tsakiris, M., Prabhu, G., & Haggard, P. (2006). Having a body versus moving your body: How agency structures body-ownership. *Consciousness and Cognition*, 15(2), 423-432.
- Tunik, E., Frey, S.H., Grafton, S.T. (2005). Virtual lesions of the anterior intraparietal area disrupt goal-dependent on-line adjustments of grasp. *Nature Neuroscience*, 8(4), 505–11.
- Uzzell, D., & Horne, N. (2006). The influence of biological sex, sexuality and gender role on interpersonal distance. *British Journal of Social Psychology*, 45(3), 579-597.
- Vainiger, D., Labruna, L., Ivry, R. B., & Lavidor, M. (2014). Beyond words: evidence for automatic language–gesture integration of symbolic gestures but not dynamic landscapes. *Psychological research*, 78, 55-69.
- Vala, J., Pereira, C. R., Lima, M. E. O., & Leyens, J. P. (2012). Intergroup time bias and racialized social relations. *Personality and Social Psychology Bulletin*, 38(4), 491-504.
- Valenti, S. S., & Gold, J. M. M. (1991). Social Affordances and Interaction I: Introduction. *Ecological Psychology*, 3(2), 77–98.
- Varela, F. J., Thompson, E., and Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
- Verfaillie, K. and Daems, A. (2002). Representing and anticipating human actions in vision. *Vis. Cogn.* 9, 217–232.
- Vesper, C., Butterfill, S., Knoblich, G., & Sebanz, N. (2010). A minimal architecture for joint action. *Neural Networks*, 23(8-9), 998-1003.
- Vigliocco, G., Kousta, S.-T., Della Rosa, P. A., Vinson, D. P., Tettamanti, M., Devlin, J. T., & Cappa, S. F. (2014). The neural representation of abstract words: The role of emotion. *Cerebral Cortex* (New York, N.Y.: 1991), 24(7), 1767–1777. <https://doi.org/10.1093/cercor/bht025>
- Villani, C., Lugli, L., Liuzza, M. T., & Borghi, A. M. (2019). Varieties of abstract concepts and their multiple dimensions. *Language and Cognition*, 11(3), 403-430.
- Villani, C., Lugli, L., Liuzza, M. T., Nicoletti, R., & Borghi, A. M. (2021). Sensorimotor and interoceptive dimensions in concrete and abstract concepts. *Journal of Memory and Language*, 116, 104173.

- Villani, C., Orsoni, M., Lugli, L., Benassi, M., & Borghi, A. M. (2022). Abstract and concrete concepts in conversation. *Scientific Reports*, 12(1), 1–15.
- Viviani, P. (2002). Motor competence in the perception of dynamic events: A tutorial. In W. Prinz & B. Hommel (Eds.), *Attention and performance XIX: Common mechanisms in perception and action* (pp. 406–442). Oxford: Oxford University Press.
- Vizioli, L., Foreman, K., Rousselet, G. A., & Caldara, R. (2010). Inverting faces elicits sensitivity to race on the N170 component: A cross-cultural study. *Journal of vision*, 10(1), 15-15.
- Vorderer, P., & Kohring, M. (2013). Permanently online: A challenge for media and communication research. *International Journal of Communication*, 7, 188–196
- Voss, M., Ingram, J. N., Haggard, P., & Wolpert, D. M. (2006). Sensorimotor attenuation by central motor command signals in the absence of movement. *Nature Neuroscience*, 9(1), 26-27.
- Walsh, L.D., Moseley, G.L., Taylor, J.L. & Gandevia, S.C. (2011). Proprioceptive signals contribute to the sense of body ownership. *The Journal of Physiology*, 589: pp. 3009–3021.
- Walther, J. B. (1996). Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research*, 23(1), 3–43. <https://doi.org/10.1177/009365096023001001>
- Walther, J. B., & Tidwell, L. C. (1995). Nonverbal cues in computer-mediated communication, and the effect of chronemics on relational communication. *Journal of Organizational Computing*, 5(4), 355–378.
- Walthier, J. B., & D’Addario, K. P. (2001). The impacts of emoticons on message interpretation in computer-mediated communication. *Social Science Computer Review*, 19(3), 324–347.
- Wang, Y., & Hamilton, A. F. D. C. (2012). Social top-down response modulation (STORM): a model of the control of mimicry in social interaction. *Frontiers in human neuroscience*, 6, 153.
- Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain*, 107(3), 829–853.
- Wassermann, E. M. (1998). Risk and safety of repetitive transcranial magnetic stimulation: report and suggested guidelines from the International Workshop on the Safety of Repetitive Transcranial Magnetic Stimulation, June 5–7, 1996. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, 108(1), 1-16.
- Watson, R., & De Gelder, B. (2017). How white and black bodies are perceived depends on what emotion is expressed. *Scientific reports*, 7(1), 41349.
- Wauters, L. N., Tellings, A. E., Van Bon, W. H., & Van Haaften, A. W. (2003). Mode of acquisition of word meanings: The viability of a theoretical construct. *Applied Psycholinguistics*, 24(3), 385-406.
- Wauters, L. N., Tellings, A. E., Van Bon, W. H., & Van Haaften, A. W. (2003). Mode of acquisition of word meanings: The viability of a theoretical construct. *Applied Psycholinguistics*, 24(3), 385–406.
- Wei, R., & Lo, V.-H. (2006). Staying connected while on the move: Cell phone use and social connectedness. *New Media & Society*, 8(1), 53–72. <https://doi.org/10.1177/1461444806059870>
- Wellman, B., Quan-Haase, A., Boase, J., Chen, W., Hampton, K., Díaz, I., & Miyata, K. (2003). The Social Affordances of the Internet for Networked Individualism. *Journal of Computer-Mediated Communication*, 8(3), 0–0.
- White, R. C., Aimola Davies, A. M., & Davies, M. (2011). Two hands are better than one: A new assessment method and a new interpretation of the non-visual illusion of self-touch. *Consciousness and Cognition*, 20(3), 956–964.
- White, R. C., Davies, A. M. A., Halleen, T. J., & Davies, M. (2010). Tactile expectations and the perception of self-touch: An investigation using the rubber hand paradigm. *Consciousness and Cognition*, 19(2), 505–519.
- Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in My insula: the common neural basis of seeing and feeling disgust. *Neuron*, 40(3), 655-664.
- Wickham, H. et al. (2019). “Welcome to the tidyverse.”. *Journal of Open Source Software*, 4(43), 1686.
- Wilhelm, F. H., Kochar, A. S., Roth, W. T., & Gross, J. J. (2001). Social anxiety and response to touch: incongruence between self-evaluative and physiological reactions. *Biological psychology*, 58(3), 181–202.
- Wilke C (2022). `ggridge`: Ridgeline Plots in 'ggplot2'. R package version 0.5.4, <<https://CRAN.R-project.org/package=ggridge>>. Willis, F. N. (1966). Initial speaking distance as a function of the speakers’ relationship. *Psychon. Sci.* 5, 221–222.
- Wilson, J., O’leary, M., Leary, O., Metiu, A., & Jett, Q. (2005). Subjective Distance in Teams. In INSEAD Working Paper Series.
- Wilson, M. and Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychol. Bull.* 131, 460–473 26 .
- Wiltermuth, S. S. (2012). Synchronous activity boosts compliance with requests to aggress. *Journal of Experimental Social Psychology*, 48(1), 453-456.
- Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological science*, 20(1), 1-5.
- Wobbrock, J. O., Findlater, L., Gergle, D., & Higgins, J. J. (2011). The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only ANOVA Procedures. In Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (pp. 143-146).

- Wohlschläger, A., Gattis, M., & Bekkering, H. (2003). Action generation and action perception in imitation: an instance of the ideomotor principle. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1431), 501-515.
- Wokke, M. E., Knot, S. L., Fouad, A., & Ridderinkhof, K. R. (2016). Conflict in the kitchen: Contextual modulation of responsiveness to affordances. *Consciousness and cognition*, 40, 141-146.
- Wolpert, D. M., Doya, K., and Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 358, 593–602.
- Wolpert, D. M., Gharamani, Z., and Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science* 269, 1880–1882. doi: 10.1126/science.7569931.
- Won, A. S., Shriram, K., & Tamir, D. I. (2018). Social Distance Increases Perceived Physical Distance. *Social Psychological and Personality Science*, 9(3), 372–380.
- Xu, J., Gannon, P. J., Emmorey, K., Smith, J. F., & Braun, A. R. (2009). Symbolic gestures and spoken language are processed by a common neural system. *Proceedings of the National Academy of Sciences*, 106(49), 20664-20669.
- Yabar, Y., Johnston, L., Miles, L., & Peace, V. (2006). Implicit behavioral mimicry: Investigating the impact of group membership. *Journal of Nonverbal Behavior*, 30, 97-113.
- Yazar, Y., Bergström, Z.M., Simons, J.S. (2017). Reduced multimodal integration of memory features following continuous theta burst stimulation of angular gyrus. *Brain Stimulation*, 10(3), 624–9.
- Zacks, J., and Michelon, P. (2005). Transformations of visuospatial images. *Behav. Cogn. Neurosci. Rev.* 4, 96–118.
- Zaferanieh, E. (2012). Conversation analysis of online chat. In *The First Conference on Language Learning & Teaching: An Interdisciplinary Approach* (pp. 30-31).
- Zeman, A., Dewar, M., & Della Sala, S. (2015). Lives without imagery—Congenital aphantasia. *Cortex*, pp. 73, 378–380.
- Zeman, A., Milton, F., Della Sala, S., Dewar, M., Frayling, T., Gaddum, J., ... & Winlove, C. (2020). Phantasia—the psychological significance of lifelong visual imagery vividness extremes. *Cortex*, 130, 426–440.
- Zhang, L. (2017). A fair game? Racial bias and repeated interaction between NBA coaches and players. *Administrative Science Quarterly*, 62(4), 603-625.
- Zhou, Y., Gao, T., Zhang, T., Li, W., Wu, T., Han, X., & Han, S. (2020). Neural dynamics of racial categorization predicts racial bias in face recognition and altruism. *Nature human behaviour*, 4(1), 69-87.
- Zhou, Z., & Geng, J. J. (2024). Learned associations serve as target proxies during difficult but not easy visual search. *Cognition*, 242, 105648.
- Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological science*, 13(2), 168–171.