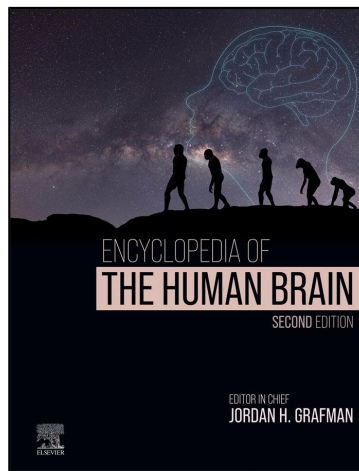


Provided for non-commercial research and educational use.
Not for reproduction, distribution or commercial use.

This chapter was originally published in Encyclopedia of the Human Brain, Second Edition, published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use, including without limitation, use in instruction at your institution, sending it to specific colleagues who you know, and providing a copy to your institution's administrator.



All other uses, reproduction and distribution, including without limitation, commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

<https://www.elsevier.com/about/policies/copyright/permissions>

Poyo Solanas, M., de Gelder, B., 2025. Social interaction—Recent behavioral and brain studies. In: Grafman, J.H. (Ed.), Encyclopedia of the Human Brain, Second Edition, vol. 4, pp. 410–421. USA: Elsevier.

ISBN: 9780128204801

Copyright © 2025 Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

Social interaction—Recent behavioral and brain studies

Marta Poyo Solanas and Beatrice de Gelder, Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, the Netherlands

© 2025 Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

Introduction	411
Understanding social interactions	411
High-level cognition	412
Importance of visual information in social interaction processing	413
Mid-level visual features	413
High-level visual features	414
Considerations in the interpretation of social interaction findings	415
Effect of task and stimulus type	415
Correct area definition	416
Hemispheric dominance	416
New directions in the study of social interactions	416
First-person perspective	416
Interaction areas or network hubs	417
Computational models to discover social interaction features	417
Concluding remarks	418
Acknowledgments	418
Conflict of interest	418
References	418

Key points

- Understanding social interactions may rely on midlevel feature computations of body expressions and social cues, rather than being driven solely by specific high-level computations.
- Biological computations involved in processing body posture and movements during social interactions may utilize midlevel features that exploit ethological characteristics of organism-environment interactions.
- Virtual reality constitutes a promising tool for the understanding of social interaction processes from a first-person perspective.
- Gaining further understanding of the underlying mechanisms of social interaction perception can contribute to unraveling the origins of psychological and neurological disorders characterized by impaired social skills, potentially guiding more effective treatment approaches.
- Social neuroscience can gain further insights on the mechanisms underlying social interaction perception by adopting computer vision approaches.

Abstract

For social species, human or non-human alike, much of their core activities take place in the context of interacting with others. The competences required to engage successfully in social interactions have only recently become the focus of neuroscientific studies. This article deals with recent studies on social interaction in dyads that have used behavioral and brain imaging methods. Initially, we examine previous explanations rooted in higher-order mentalistic concepts. Subsequently, we delve into more contemporary approaches that focus on the visual processes involved in social interaction, including the identification of specific brain areas and networks dedicated to processing such interactions. Furthermore, we explore the potential of virtual reality for investigating realistic social interaction scenarios. Methodological challenges associated with processing complex visual information from naturalistic interaction scenes are also addressed, along with an overview of computational approaches aimed at managing this complexity.

Introduction

Engagement in social interactions, whether as passive observers or as active participants, is a substantial part of daily life for many species. Real life social interactions mobilize various skills related to perceiving and understanding social signals provided by conspecifics, along with skills for preparing and executing appropriate and adaptive reactions. On the perception side, interacting with others in naturalistic circumstances often entails the integration of multiple senses, combining visual cues (e.g., face, body), auditory signals (e.g., voice) and other sensations (e.g., interoception, touch). On the cognitive side, interaction relies on competences required for understanding rules of interactive logic, encompassing various social phenomena such as altruism and empathy (Decety and Jackson, 2004).

In view of the centrality of social interactions in daily life, it is no surprise that many psychiatric disorders exhibit problems related to this strong social component (Konrad et al., 2024). Several psychopathologies are associated with difficulties in social interaction, with conditions such as Autism Spectrum Disorder (Tager-Flusberg, 2007) and schizophrenia (Lavelle et al., 2014) being particularly notable due to the pronounced symptomatic nature of impaired social engagement. Moreover, there is a growing awareness that in the broader population, exclusion from social interaction and social ostracism can lead to subjective experiences akin to physical pain (Eisenberger, 2012).

It is therefore surprising to note that for behavioral neuroscience, the study of the brain basis of social interaction is a relatively new field of research. As we will see, the literature is still very limited and only in the last couple of years did we witness explicit arguments in favor of a novel orientation of neuroscience towards social interaction (Hari and Kujala, 2009). This is not to say that social phenomena have not already been extensively investigated in the past decades. In fact, the notion of a social brain has been present in the human neuroscience literature for quite a while and studies on the neural basis of social behavior broadly defined are numerous (Brothers et al., 1990). A recent meta-analysis reported over 5000 publications related to social behavior studies using brain-imaging methods (Feng et al., 2021). The topics included were diverse, ranging from norm violation to empathy and mentalizing, to name a few. Although important details concerning major differences in stimuli and tasks were sidestepped, the study concluded that four major brain networks are involved in social processes: the salience network, the subcortical network, the default mode network and the executive control network (Feng et al., 2021).

While many social phenomena have been investigated in the past decades, as indicated above, not all aspects relate genuinely to social interactions. This is particularly evident in phenomena like empathy or joint action, where the actions of one individual may not necessarily influence the other. In contrast, some events and actions are inherently related to social interactions, as their primary objective is to influence others and achieve a change in their behavior. Given the multidimensionality of social behavior, it becomes pertinent to narrow down the definition of social interactions. Yet, even within the study of social interactions, there is a broad range of aspects to investigate.

Some interaction studies focus on scenarios where interaction is defined by a shared goal, that is, the action of each agent is defined by the effect it has on the other. Examples include acts like aggression or when someone waves at another to attract their attention. Such genuine interaction movements are associated with physical characteristics such as distance between the agents and their orientation towards each other. In line with this, available studies have focused on physical determinants (e.g., velocity, distance, direction of movement) to understand social interaction perception (e.g., Keck et al., 2022). However, these factors alone may not be sufficient to define interaction or capture its essence. As we shall discuss later, it may also matter differently depending on how one conceives interaction. For example, some actions cannot be understood if both agents are not shown simultaneously, other actions may lose their meaning when shown at a velocity below or above the normal. However, this aspect has not been yet investigated systematically.

Another issue is that the focus of most experiments has been on investigating how passive observers react to social signals, mostly in the field of visual and auditory perception. The design of such studies tend to overlook the active role often inherent in social interactions. One could contend that there remains a hidden but crucial interactive dimension in such experiments, as presenting participants with faces, bodies, or voices effectively places the observer in a virtual interaction scenario. However, this interactive aspect has largely been overlooked, given that the perception of social signals, such as facial expressions, is conventionally studied as a subset of object perception. This narrow perspective tends to obscure the interactive dimension.

Considering the multifaceted nature of social interactions, in this article we focus on the discussion of recent studies investigating social interactions with the use of point light displays (PLD; i.e., visual representations of motion using small, bright points of light, typically placed on key joints of a person's body) or naturalistic videos showing two whole body agents. Therefore, this review excludes studies of social interactions beyond dyads as well as research that just focuses on facial expressions. We review behavioral papers as well as neuroimaging research, principally using functional magnetic resonance imaging (fMRI).

Understanding social interactions

To navigate our social environment, the brain must process the rich information provided by social interactions. This not only includes the fast processing of incoming visual information but also the recognition of the mental states of our conspecifics and the prediction, whether conscious or not, of their future behavior based on that knowledge (Frith, 2007). In line with this, social interaction literature highlights neural correlates spanning from high-level visual areas to regions associated with mentalization processes ("theory of mind"; Frith and Frith, 2006). Together, they comprise a broad set of regions referred to as the social brain

(Brothers et al., 1990), often including, but not limited to, the posterior superior temporal sulcus (pSTS), and the adjacent temporo-parietal junction (TPJ), the anterior superior temporal sulcus (aSTS), the medial prefrontal cortex (MPFC) and the amygdala (Bedny et al., 2009; Frith, 2007; Gallagher and Frith, 2003; Saxe, 2006).

Traditionally, research in social psychology has tended to prioritize high-level social cognition, including complex mental processes presumably at the basis of our ability to decipher intentions conveyed through facial expressions, bodily movements and social interactions. However, the neural representation of intentions, especially social ones in interactive situations, remains a topic of debate, with questions about the involvement of dedicated brain structures (Frith, 2007). More recently, there has been a notable theoretical and empirical shift towards recognizing the pivotal role of visual cues in understanding social behaviors (de Gelder and Poyo Solanas, 2021; McMahon and Isik, 2023). This shift is crucial, as it acknowledges the brain's innate ability for rapid and automatic processing of social information and the evolutionary importance of such behaviors. In the subsequent subsections, we retrace the chronological sequence of social neuroscience research. Initially, we delve into the high-level cognitive processes implicated in social interaction processing before shifting our attention to the representation of social interaction cues within the visual system and exploring their behavioral processing.

High-level cognition

Anticipating others' actions and comprehending their intentions are crucial components for thriving in social environments. The extent to which these predictive abilities are governed by related or separate neural networks remains a topic of ongoing debate. Key evidence in this debate revolves around the mirror (Gallese et al., 1996) and mentalizing ("theory of mind"; Frith and Frith, 2006) networks.

The Mirror Neuron System (MNS), also known as the Action Observation Network (AON), is a network of brain areas involved in understanding others' actions and intentions (Fogassi et al., 2005; Iacoboni et al., 2005; Jellema et al., 2000; Rizzolatti et al., 2001). Found in humans and some primates, the MNS includes areas such as the pSTS, inferior parietal lobule (IPL), superior parietal lobule, and intraparietal sulcus (IPS), along with the dorsal and ventral premotor cortex and the posterior part of the inferior frontal gyrus (Caspers et al., 2010; Molenberghs et al., 2012) (Fig. 1). Among these, the pSTS has been suggested to play a role in social interaction perception from studies involving this region in processing animacy when two objects interact (Schultz et al., 2005) and when comparing the processing of two agents interacting versus two agents performing non-social movements (Centelles et al., 2011).

According to the mirror neuron theory, understanding other's actions and intentions occurs through simulating observed actions within one's own motor repertoire (Rizzolatti and Craighero, 2004; Rizzolatti and Sinigaglia, 2010). Initially identified through shared activation of brain areas during instrumental action perception, the MNS has also been implicated in processing emotions observed in others. This extension was inspired by findings of a common neural response observed when witnessing others' emotions as well as when we experience the same emotion ourselves (Wicker et al., 2003). In the context of social interactions, this poses a challenge, as the observer must map the actions of multiple conspecifics onto their own motor repertoire simultaneously. Thus far, it remains unclear whether motor simulation plays a role in understanding the social intentions arising from interactions between two individuals (Brass et al., 2007; Jacob and Jeannerod, 2005; Saxe, 2005).

On the other hand, The Mentalizing System, also known as the Theory of Mind network or the Social Cognition network, encompasses brain regions involved in understanding and interpreting both one's own and others' thoughts, feelings, and intentions (Frith and Frith, 2006). Activation of this network occurs when individuals consider mental states, regardless of attention to movements or visual cues, as observed during tasks involving narrative reading, animation viewing or false belief. This system shares overlap with the Default Mode Network (DMN), indicating a connection between representing others' mental states and self-referential processes (Smallwood et al., 2021). Key regions within this network include the medial prefrontal cortex, temporo-parietal junction, and precuneus (Saxe, 2006) (Fig. 1).

Despite the frequent coactivation of the Mirror and Mentalizing networks in individual action literature, there is compelling evidence for a functional separation between them. Various factors argue in favor of a distinction between the two, including task type (e.g., implicit versus explicit) and discerning "what" (action) versus "why" (mental state) a person is doing (Chiavarino et al., 2012; Spunt et al., 2016; Spunt and Lieberman, 2012a,b). In the case of social interactions, this division becomes more intricate given the higher complexity of these stimuli. While several studies have shown a simultaneous engagement of these networks irrespective of the type of stimuli (Arioli et al., 2018; Centelles et al., 2011; Iacoboni et al., 2004; Kujala et al., 2012), others have suggested different activation and connectivity patterns based on interaction dimensions such as cooperativity and affectivity (Arioli et al., 2018; Canessa et al., 2012; Proverbio et al., 2011). Further research is therefore necessary to understand the functional segregation as well as the shared roles of these two networks in relation to the multidimensionality of social interaction perception.

In addition to the MNS and Mentalizing system, a recent meta-analysis involved the set of regions typically associated to the Salience and the Subcortical Networks (SN and SCN, respectively) in the processing of social interactions (Feng et al., 2021). These networks form a common neural motivation system suggested to be involved in decision making in a wide range of social situations (Grabenhorst and Rolls, 2011; Gu et al., 2019; Izuma et al., 2008; Levy and Glimcher, 2012; Lin et al., 2012; Rangel et al., 2008; Saxe and Haushofer, 2008). Therefore, it is not surprising that these networks are also engaged during the processing of social interactions. However, while the SCN is mainly activated by positive social interactions and norm adherence, the SN is primarily involved in negative social interactions and norm violations (Feng et al., 2021). Furthermore, particularly in situations of norm violation during social interactions, another network, the Central Executive Network, has shown to be involved in information selection

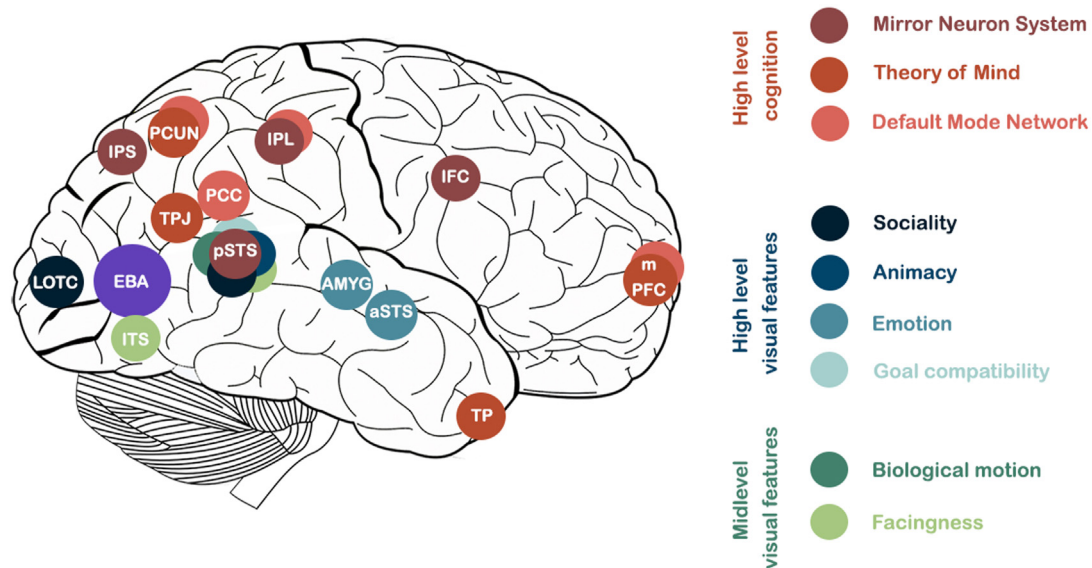


Fig. 1 Overview of areas engaged in social interaction perception. Several high-level cognitive networks are activated during social interaction perception, including the Mirror Neuron System, the Theory of Mind network, and the Default Mode Network. The Mirror Neuron System comprises areas such as the posterior superior temporal sulcus (pSTS), inferior parietal lobule (IPL), intraparietal sulcus (IPS), and the posterior part of the inferior frontal gyrus (IFG) (Caspers et al., 2010; Molenberghs et al., 2012). The Theory of Mind network includes regions like the precuneus (PCUN), temporo-parietal junction (TPJ), medial prefrontal cortex (mPFC), and temporal pole (TP) (Saxe, 2006). The Default Mode Network involves areas such as the PCUN, posterior cingulate cortex (PCC), and mPFC (Feng et al., 2021). Several areas have been associated with high-level visual feature representation during social interaction perception, such as sociality in the pSTS and LOTC (Tarhan and Konkle, 2020), animacy in the pSTS (Schultz et al., 2005), emotion/interaction type in the anterior STS (aSTS) (McMahon et al., 2023) and amygdala (Sinke et al., 2010), and goal compatibility in the pSTS (Isik et al., 2017; Walbrin et al., 2018). The extrastriate body area (EBA) is selective to bodies, also in dyads (Abassi and Papeo, 2020). Several areas represent midlevel visual features related to social interactions, including the pSTS for biological motion (Grossman et al., 2000) and the inferior temporal sulcus (ITS), EBA, and pSTS for facingness (Bellot et al., 2021).

and integration for effective decision-making (Buckholz and Meyer-Lindenberg, 2012; Feng et al., 2021; Miller and Cohen, 2001) (Fig. 1).

Importance of visual information in social interaction processing

So far, we have seen that social interaction perception has traditionally been associated with higher-level social cognition. However, emerging evidence suggests that recognizing social interactions may rely on fast, automatic, and visually driven processes rather than solely on complex mental models (Baker et al., 2022; McMahon and Isik, 2023). This is supported by studies showing that social interactions are given precedence in visual search tasks (Papeo et al., 2019) as well as in accessing conscious awareness during binocular rivalry tasks (Su et al., 2016). Even the pioneering research conducted by Heider and Simmel (1944) demonstrated that humans possess the ability to discern intricate details about others' interactions based solely on basic visual cues. Interestingly, Pitcher and Ungerleider (2021) have recently proposed a third visual pathway, alongside the traditional dorsal and ventral streams, suggested to support social perception. This newly proposed pathway spans from the early visual cortex (EVC) to the superior temporal sulcus (STS), with representations of increasingly complex features (Pitcher and Ungerleider, 2021). In the following subsections, we will see how this proposal is further supported by recent findings showing rich representations of social interactions within the visual system beyond basic visual cues (Baker et al., 2022; McMahon and Isik, 2023).

Mid-level visual features

Recent research in social and affective perception has highlighted the significance of body postures and movements in understanding social behaviors and interactions (de Gelder and Poyo Solanas, 2021, 2022). The importance of these bodily cues, termed mid-level body features, resides in the fact that they bridge the gap between low-level visual attributes (e.g., motion energy) and high-level concepts (e.g., valence or emotion expression) while still containing rich information (de Gelder and Poyo Solanas, 2021, 2022). This line of research has already shown that kinematic features such as velocity, acceleration and body part displacement are crucial for recognizing emotional body expressions (Atkinson et al., 2007). Postural body features have also proved to be essential for accurately recognizing emotional body expressions, such as limb contraction in the perception of fearful body movements or the directionality of movement for distinguishing angry from fearful expressions (Poyo Solanas et al., 2020a,b).

Recent advances in neuroscience have taken a step forward and begun to explore the brain regions responsible for encoding mid-level visual body cues associated with social actions, with a specific focus on the visual cortex. It is hypothesized that these

representations are likely formed through hierarchical computations of visuospatial and motion cues (McMahon and Isik, 2023), particularly in areas such as the extrastriate body area (EBA), in the lateral occipitotemporal cortex (LOTC), and the pSTS. EBA demonstrates selective responses to bodies, particularly in dyads (Abassi and Papeo, 2020), and exhibits features of configural processing in dynamic displays of point light figures (Abassi and Papeo, 2022). In a recent study, Poyo Solanas et al. (2020b) have linked the pSTS to discerning fear expressions from other emotional expressions through the analysis of kinematic body cues, thus highlighting its crucial role in social cognition. In line with this, this region has also been implicated in various social functions such as biological motion processing, speech processing, audiovisual integration, and theory of mind (Deen et al., 2015; Grossman et al., 2000; Hein and Knight, 2008), and has shown responses to social interactions in both controlled (Isik et al., 2017; Walbrin et al., 2018) and naturalistic stimuli (Masson and Isik, 2021).

While single-person features, such as the postural and kinematic body features reported above, are important for understanding social behaviors, they still do not capture the interactive nature of social interactions entirely. Recent research has thus started to emphasize the importance of features capturing interactive behaviors when investigating social interaction perception. For instance, research by Keck et al. (2022) demonstrated that individuals rely on a combination of intra (e.g., velocity, limb angles) and interpersonal features (e.g., interpersonal distance, synchrony) when judging emotional expressions in interactive scenarios. The effectiveness of these features varied based on the specific emotions displayed, highlighting the nuanced role of these features in social perception. For instance, intrapersonal velocity mattered for some interactions, while personal distance was crucial for others (Keck et al., 2022). However, this study did not address comprehensively what constitutes an interaction, treating the set of basic emotions uniformly and overlooking differences in implicit interactivity (e.g., sadness does not directly involve an addressee, whereas anger does).

The context-dependency of one of these interpersonal features, interpersonal distance (i.e., distance between communicating individuals), was also observed in other studies, and appeared to be influenced by factors such as emotional and motivational relevance (Brennan and Martin, 2012; Iachini et al., 2014). For example, Iachini et al. (2014) discovered that peripersonal space increases in hostile or uncomfortable situations but decreases in friendly and comfortable situations. Investigating the impact of conversation type (argumentative or friendly) on body movement dynamics, Paxton and Dale (2017) found that nonverbal behavior during a conversation changes based on contextual constraints, with arguments, for instance, decreasing movement synchrony (Paxton and Dale, 2017). This, in turn, can influence the distance between individuals during an interaction.

Another prominent characteristic of social interactions is “facingness,” which refers to whether actors are oriented toward each other (Papeo, 2020). Research indicates that dyads are recognized more quickly when individuals face each other compared to non-facing dyads (Papeo et al., 2019). This phenomenon extends to groups of people as well (Vestner et al., 2019). Moreover, studies have shown that the visual system is more proficient at detecting agents when they are depicted interacting with another individual (Manera et al., 2011; Neri et al., 2006). Su et al. (2016) conducted an intriguing investigation exploring the prioritization of stimuli containing dynamic social interaction information for conscious perception over non-interacting dyads. The study compared facing versus non-facing dyads, as well as dyads with coupled movements versus non-coupled movements, using dynamic images. The findings suggested that coupled movement enhances perceptual strength compared to facing alone, leading to greater awareness (Su et al., 2016). These findings highlight the importance of investigating the relationship between different interaction features.

Studies like the one conducted by Bellot et al. (2021) have gone beyond exploring the influence of factors such as “facingness” and interpersonal distance on interaction perception to investigate their neural representations. This study involved facing and non-facing dyads of individuals performing simple actions while either moving towards or away from each other. Activation in EBA and enhanced connectivity to the pSTS were observed only when actors faced each other. Consequently, the study concluded that facing/spatial distance between actors influenced interaction perception, regardless of the action type (Bellot et al., 2021). Notably, unlike previous findings by Su et al. (2016), the study did not include coordinated or complementary actions between facing actors. The design focused solely on the aspects of moving towards/facing and spatial distance, with no manipulation of factors such as the direction of movement or the specific actions performed.

Similarly, McMahon et al. (2023) found that the EBA encodes facing direction and spatial proximity, but upon considering additional factors, it did not exhibit distinct representations of these social cues. Instead, unique variance explained by these features was observed in the nearby lateral occipital cortex. This study is of particular interest as it uncovered a trend of increasing feature complexity within the third visual pathway. Specifically, it found that low-level attributes are predominantly encoded in early visual cortex and MT, while scene, object, and basic social features find representation in the LOC, and social interaction features are processed along the STS (McMahon et al., 2023). Similar gradient of feature representations across the visual cortex hierarchy was also reported by Masson and Isik (2021).

High-level visual features

In addition to midlevel social cues, it has been shown that the visual cortex also encodes higher-level visual features of social behaviors and interactions. For example, recent research suggests that sociality is an important organizing dimension of action representations in both behavior and brain function, particularly in lateral occipital visual regions (Tarhan and Konkle, 2020). It has also been reported that the EBA and pSTS are able to differentiate between dyadic interactions types (Walbrin and Koldewyn, 2019). Furthermore, pSTS has been shown to play a role in the processing of goal compatibility between agents, such as whether an interaction involves helping versus hindering or cooperation versus competition (Isik et al., 2017; Walbrin et al., 2018). Recent research further suggests that the STS, especially anterior regions, may encode affective information in social interactions, distinguishing between positive and negative interactions (McMahon et al., 2023). However, despite these findings, affective features do not

seem to account for unique variance along the lateral stream (McMahon et al., 2023). Additionally, McMahon et al. (2023) indicated that STS responses to social interactions are driven by communicative actions that are not limited to one's social partners but also in broader contexts. Yet, it must be noted that similar results have also been observed in the case of the TPJ within the theory of mind network. As McMahon and Isik (2023) pointed out, the contribution of visual input versus top-down signals from the theory of mind network in processing interaction goals remains unclear given the low temporal resolution of fMRI (McMahon and Isik, 2023).

Overall, recent research underscores the involvement of the visual cortex in detecting social interactions, suggesting a dedicated visual pathway for social perception (McMahon et al., 2023; Pitcher and Ungerleider, 2021; Wurm and Caramazza, 2022). While the distinction between recognizing social interactions and theory of mind processes is evident, further research is needed to elucidate their neural mechanisms and their interplay during naturalistic viewing. A comprehensive picture would be one that considers features of social interaction at the different levels at which naturalistic videos can be described and the relation of these feature results to brain processes.

Considerations in the interpretation of social interaction findings

Considerable evidence suggests a functional separation between action observation and mentalizing networks when processing individual actions based on factors such as the type of information processed, task demands, and the distinction between behavioral and mental states (Chiavarino et al., 2012; Spunt et al., 2016; Spunt and Lieberman, 2012a,b). However, this division appears to be challenged when processing social interactions, as both networks are concurrently activated regardless of the type of stimuli used (Arioli et al., 2018; Centelles et al., 2011; Iacoboni et al., 2004; Kujala et al., 2012). This simultaneous engagement may signify the increased complexity involved in understanding social interactions. To ensure accurate interpretation of findings, it is crucial to control for confounding factors such as task type, stimuli, brain hemispheric dominance, and precise area definition. These considerations extend beyond the differentiation of action observation and mentalizing networks, as these factors have demonstrated an impact on the visual representation of social stimuli. Therefore, maintaining careful control and awareness of these factors in experimental designs is essential to prevent misinterpretation of neural activation patterns and safeguard the validity of research results. The subsequent subsections will delve into these confounding elements within the context of recent research.

Effect of task and stimulus type

One key aspect when investigating social interactions is the type of interaction depicted, as different interactions may engage distinct neural substrates. For example, the study conducted by Sinke et al. (2010) exemplifies the importance of emotional content, as well as task conditions, in shaping brain activity patterns during social interactions. In this study, participants were exposed to naturalistic videos portraying teasing or threatening interactions while either engaging in an emotion-related task (i.e., emotional categorization of interaction type) or in a non-emotional task (i.e., color discrimination of fixation dot) in a carefully balanced task design. The objective was to identify differences in brain activity when participants viewed the videos under different task conditions and to ascertain whether body-sensitive regions showed differences between conditions and tasks (Sinke et al., 2010). Higher activation was observed in regions like the amygdala during threatening interactions, with distinct patterns emerging between the emotion and color naming tasks. Intriguingly, body-sensitive areas such as the EBA and the STS exhibited increased activity during threatening interactions when participants engaged in the color-discrimination task (Sinke et al., 2010), suggesting automatic processing of emotional cues irrespective of explicit attention.

Certainly, the role of attention introduces an additional layer of complexity to the neural processing of social interactions, especially in situations where interactions exhibit asymmetry, such as during a fight or a burglary. In such instances, observers typically do not attend to both participants simultaneously; instead, they alternate their focus between the two, concentrating on either the victim or the aggressor based on the nature of the interaction. A study by Van den Stock et al. (2015) delved into this aspect by employing realistic video clips portraying aggressive two-person interactions, distinctly identifying one individual as the aggressor and the other as the victim. To manipulate attention, the study incorporated variations in attentional focus (on either the aggressor or the victim) and task demands (identifying the type of interaction or discerning the color of a dot displayed in the videos). When attention was directed towards the aggressor, heightened activity was observed in brain regions including the amygdala, EBA, anterior cingulate cortex (ACC), STS, and the occipital pole (Van den Stock et al., 2015). However, no activity differences specific to the victim-focused condition were observed, underscoring the selective impact of attention on neural responses during the observation of asymmetric social interactions.

Apart from the type of stimuli, tasks may engage both perceptual and mentalization systems, emphasizing the importance of using appropriate tasks or analyses to differentiate them effectively. For example, a task requiring button presses might result in the engagement of the prefrontal cortex, which could be misattributed to mentalizing activity (Centelles et al., 2011; Sapey-Triomphe et al., 2017). Moreover, the use of an active versus a passive task where only attention is controlled for may influence the results. For instance, Landsiedel et al. (2022) employed a color detection task and found that no activity in brain regions associated with social interaction processing. This is consistent with results from a study by Marrazzo et al. (2021) reporting that when body perception was not the focus of the task, activity in the category sensitive areas decreased to the extent of becoming non-significant. However, these results contrast with findings from other studies (e.g., Sinke et al., 2010). Therefore, careful consideration

of experimental designs is necessary to avoid task confounds in the investigation of social interaction and mentalization processes (Centelles et al., 2011; Sapey-Triomphe et al., 2017).

In addition to the type of task or the meaning of the interaction, the type of stimulus, whether static or dynamic, significantly influences the investigation of social interaction processing in the brain. Initially, studies predominantly employed static images of generic humans, but subsequent research highlighted the importance of movement in delineating neural responses to interactions. Landsiedel et al. (2022) recently addressed this issue by comparing brain activity in response to naturalistic stimuli, point light displays, and static images. Their findings underscored the differential neural responses elicited by PLD compared to full naturalistic stimuli, emphasizing the critical role of dynamic information in understanding the neural mechanisms underlying social interactions (Landsiedel et al., 2022).

Correct area definition

A crucial aspect to consider is the spatial proximity of brain regions that may serve distinct roles or belong to different networks. Take, for example, the pSTS and TPJ, neighboring regions within the temporal cortex. While both regions contribute to social cognition, the pSTS is more focused on perceptual aspects of social information processing, whereas the TPJ is more involved in higher-level cognitive processes related to understanding the mental states of others (Bahnmann et al., 2010). Similarly, in the realm of single-body research, multiple studies have indicated that the EBA is not a uniform area but comprises at least three distinct and relatively autonomous regions. This is supported by findings regarding functional stimulus comparisons, visual field maps as well as anatomical landmarks (Weiner and Grill-Spector, 2011). Incorrect identification of these areas might lead to a misinterpretation of their roles.

Hemispheric dominance

It is currently not clear whether there exists a clear pattern of laterality for social interaction perception. Recent research utilizing transcranial magnetic stimulation (TMS) in combination with fMRI indicated a prominent dominance of the left hemisphere (LH) (Gandolfo et al., 2024). Particularly, this study utilized dyadic stimuli to explore how the brain processes facing and non-facing dyads, revealing a lateralized response in the EBA during fMRI. This unexpected finding was further supported by TMS targeting the left EBA, which effectively eliminated the inversion effect observed for facing dyads (Gandolfo et al., 2024). A hemispheric division of the TPJ has also been proposed: while the TPJ is crucial for understanding “private intentions,” the left TPJ may have a specific function in comprehending social intentions (Ciaramidaro et al., 2007). Although these studies opened avenues for further exploration into the neural basis of social interaction perception, further investigation is needed to determine whether a consistent pattern of laterality exists for social interaction perception.

In conclusion, when these considerations are addressed through careful experimental design, such as using within-subjects designs without task confounds and employing stimuli or analysis methods that disentangle social interaction recognition from theory of mind (McMahon and Isik, 2023), distinct neural networks underlying social interactions and mentalization may emerge.

New directions in the study of social interactions

First-person perspective

Traditionally, studies focusing on face, body or interaction processing assume that the experimental participant is passive in the sense of having only the role of a third person observer. However, when observing interactions involving conflict or confrontation, the observer's brain may be triggered into actively taking a position, reacting like a virtual participant. Measuring such potential observer engagement is central in understanding the dynamics of social interaction. In this regard, recent technological advancements offer new avenues for studying participant involvement in social interactions, moving beyond the traditional third-person perspective. One of these recent innovative approaches involves implementing social interaction through embodiment in Virtual Reality (VR). Particularly, VR enables researchers to immerse participants in realistic interaction scenarios, allowing for investigations into the participant's experience of interacting with others (Sanchez-Vives and Slater, 2005). This is further enhanced by incorporating additional body tracking technologies allowing participants to synchronize the movements of their virtual body with their real-time body movements, thereby establishing a sense of agency and visuo-motor synchrony between the real and artificial body (Kokkinara et al., 2015).

The importance of VR for social interaction research has been supported by studies demonstrating its ability to induce perceptual illusions of presence and plausibility, prompting participants to exhibit realistic behaviors despite being aware of the virtual nature of simulations (Slater, 2009). Regarding first-person perspective, this has been exemplified in the study by de Borst et al. (2020), where fMRI data were acquired while participants either embodied a female victim of verbal domestic violence from a first-person perspective (1PP) or observed the situation from a third-person perspective (3PP). The results indicated enhanced synchronization of fronto-parietal brain networks and amygdala activity when threats were perceived as directed towards oneself (1PP) compared to another person (3PP) (de Borst et al., 2020). These findings underscore the significance of brain areas involved in encoding aspects of bodily self and threat perception. Moreover, a higher sense of body ownership and self-identification with the victim was observed in 1PP compared to 3PP, aligning with findings by Gonzalez-Liencre et al. (2020). Therefore, the

investigation of first-person perspective processes in the study of social interactions proves crucial, shedding light on nuanced aspects of embodied experiences on our understanding of social dynamics.

In addition to exploring the neural underpinnings of social interactions from a 1PP, VR provides an opportunity to examine physiological responses within social contexts. For instance, [Mello et al. \(2022\)](#) combined VR with heart rate monitoring and a power platform to track physical movement in three-dimensional space. Their findings revealed decreased heart rate and diminished postural mobility among participants when confronted with aggressive and closely positioned virtual characters ([Mello et al., 2022](#)). These combined methodologies present promising opportunities for gaining deeper insights into the dynamics of social interaction within controlled and ecologically valid environments.

Moreover, it has been shown that VR can induce measurable behavioral and cognitive effects in the participants. This is exemplified in the study by [Seinfeld et al. \(2021\)](#), who investigated the brain mechanisms underlying enhanced emotion recognition after embodying a domestic violence victim in VR. Participants underwent two fMRI scanning sessions, once before and once after experiencing a VR scenario, alongside pre- and post-behavioral assessments of emotional facial categorization. Findings revealed that embodying a female victim of domestic abuse not only heightened the recognition of fearful female faces but also implicated the Default Mode Network in these alterations. Specifically, there was augmented DMN activity when processing ambiguous emotional expressions following the VR experience, juxtaposed with reduced activity distinctly for fearful expressions ([Seinfeld et al., 2021](#)). Therefore, these findings underscore the potential of VR in enhancing our understanding of social dynamics and their profound impact on behavior and cognition.

Interaction areas or network hubs

Adopting the rationale of many studies on face and body selectivity in the brain, researchers have set out to investigate the brain basis of social interaction in the hope of identifying a dedicated brain area ([Abassi and Papeo, 2024](#)). The central hypothesis is that a brain area can be found that is selectively sensitive to social interactions. Studies on category selectivity suggest that certain brain regions such as EBA, fusiform body area (FBA) and pSTS play a role, but also that their roles vary depending on factors like the type of interaction, what participants attend to, the task settings and last but not least, the participant psychology ([de Gelder and Poyo Solanas, 2021](#)). However, understanding social interaction selectivity involves considering a broader network of brain areas, rather than focusing solely on one or another specific region ([Feng et al., 2021](#)). Research findings on mentalizing and mirror neuron networks have already suggested that interaction selectivity involves various brain areas working together ([Centelles et al., 2011](#)). Therefore, promising approaches for investigating social interactions must involve identifying one or more networks sustaining interaction perception and exploring how its activity and connectivity with other brain areas support specific interactions such as, for example, defensive or aggressive behaviors. To achieve this, it is essential to delve into the specific computational mechanisms involved, how connectivity contributes to interaction sensitivity, and the correlation between this sensitivity and broader body processing, as evidenced in numerous prior studies. This approach contrasts with the traditional method of isolating interaction-selective areas first and then mapping their connections. Instead, it suggests studying interactions at a network level, similar to the approach taken in whole-body processing research ([de Gelder and Poyo Solanas, 2021](#); [Li et al., 2023, 2024](#); [Weiner and Grill-Spector, 2011](#)).

Computational models to discover social interaction features

Investigating social interactions without a structured approach can pose several challenges due to the inherent complexity. Traditional methods have typically relied on observational approaches and face challenges such as dimensionality and the subjective interpretations of observers. In this regard, computational approaches offer a possible solution by reducing complexity and identifying key features from the multitude of available data. So far, most models aiming at replicating human social interaction recognition rely on explicit mental representations of the meaning of the interaction and the mental state of the agent ([Kiley Hamlin et al., 2013](#); [Shu et al., 2021](#); [Ullman et al., 2009](#)). Known as generative inverse planning models, these top-down approaches tend to be computationally intensive as they involve simulating multiple possible behaviors given different internal states and environmental conditions. Their objective is to deduce the underlying intentions, goals, or mental states of agents by scrutinizing their observed behaviors or actions within a specific environment. However, recent studies suggest that bottom-up discriminative models with proper inductive biases can also explain human social interaction recognition in a less computationally intensive manner ([McMahon and Isik, 2023](#)). Yet, and despite neuroscientific evidence also highlighting the importance of visual social primitives, effective bottom-up visual models for recognizing social interactions remain scarce ([McMahon and Isik, 2023](#)).

A novel model for recognizing social interactions, called SocialGNN, has been proposed recently, emphasizing the significance of visual primitives ([Malik and Isik, 2023](#)). SocialGNN utilizes graph neural networks to incorporate relational dynamics among interacting agents. Notably, this model outperforms others in its alignment with human assessments. However, the critical question remains: do these identified features align with how the human brain processes visual information? Developing a biologically plausible model necessitates mirroring the functions of the visual cortex and other relevant brain regions involved in perceiving body language ([Giese and Poggio, 2003](#)), a task that is still in progress. Adopting a neuroethological framework may guide the model's development and assist in formulating testable hypotheses about the essential elements of social interaction perception ([de Gelder and Poyo Solanas, 2021](#); [Hasson et al., 2020](#)).

Concluding remarks

Social neuroscience is undergoing a shift in its approach to studying social processes (de Gelder and Poyo Solanas, 2021; McMahon and Isik, 2023). Traditionally, research focused on higher-level aspects of social cognition, like mentalization processes for understanding body expressions and interactions (Frith and Frith, 2006). However, recent evidence suggests that visual cues play a fundamental role in recognizing social behaviors and interactions (de Gelder and Poyo Solanas, 2021; McMahon and Isik, 2023). Multiple brain areas, including the right pSTS and TPJ, are involved in various social cognitive processes such as perceiving biological motion, mentalizing, and making moral judgments. Moving beyond a high-level approach and separating between visual and mentalizing processes in the study of social interactions can shed light to the specific contributions of the various areas involved. For example, it has already been put forward a two-stage model where the right pSTS initially breaks down visuospatial information into discrete elements. The TPJ then engages in more complex computations related to interpreting others' behavior, suggesting a hierarchical increase in activity (Bahnemann et al., 2010). Recent connectivity evidence supports this model, indicating the pSTS has a crucial role in social interaction perception, via both bottom-up visuomotor processing and top-down attribution of affective/mental states, influencing both action observation and mentalizing networks (Arioli et al., 2018).

This shift promises a novel understanding of the visual mechanisms underlying social cognition. However, many questions and challenges remain, particularly regarding characterizing the computational goal of social perception (McMahon and Isik, 2023). Social interactions may require more flexible representations compared to object recognition (de la Rosa et al., 2014), and further research is needed to understand these neural computations and how they integrate with other perceptual modalities and cognitive systems (McMahon and Isik, 2023). In this regard, focusing on midlevel features has significant implications for developmental and clinical studies. Understanding how midlevel features contribute to the development of body-selective areas in the brain can provide insights into nonverbal emotion deficits and inform treatment options for disorders like Autism Spectrum Disorder. Integrating findings from midlevel feature research may help bridge gaps between high-level cognitive theories and different levels of visual processing (de Gelder and Poyo Solanas, 2021), such as theories regarding impaired biological movement perception (Blake et al., 2003). This integration could facilitate the development of more precise interventions.

In conclusion, investigations on midlevel features offer a promising avenue for understanding rapid expression perception and action preparation in social communication. Further research is needed to develop a comprehensive theoretical framework and computational models for understanding midlevel vision processes sustaining social interaction. This approach promises to advance substantially our understanding of organism-environment interaction and the neural mechanisms underlying social cognition.

Acknowledgments

This work was supported by the ERC Synergy grant (Grant number: 856495; Relevance), by a Horizon 2020 grant (Grant number: 101070278; ReSilence), and by a Horizon 2020 grant (Grant number: 101017884; GuestXR).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Abassi, E., Papeo, L., 2020. The representation of two-body shapes in the human visual cortex. *J. Neurosci.* 40, 852–863. <https://doi.org/10.1523/JNEUROSCI.1378-19.2019>.
- Abassi, E., Papeo, L., 2022. Behavioral and neural markers of visual configural processing in social scene perception. *Neuroimage* 260, 119506. <https://doi.org/10.1016/j.neuroimage.2022.119506>.
- Abassi, E., Papeo, L., 2024. Category-selective representation of relationships in the visual cortex. *J. Neurosci.* 44, e0250232023. <https://doi.org/10.1523/JNEUROSCI.0250-23.2023>.
- Arioli, M., Perani, D., Cappa, S., Proverbio, A.M., Zani, A., Falini, A., Canessa, N., 2018. Affective and cooperative social interactions modulate effective connectivity within and between the mirror and mentalizing systems. *Hum. Brain Mapp.* 39, 1412–1427. <https://doi.org/10.1002/hbm.23930>.
- Atkinson, A.P., Tunstall, M.L., Dittrich, W.H., 2007. Evidence for distinct contributions of form and motion information to the recognition of emotions from body gestures. *Cognition* 104, 59–72. <https://doi.org/10.1016/j.cognition.2006.05.005>.
- Bahnemann, M., Dziobek, I., Prehn, K., Wolf, I., Heekeren, H.R., 2010. Sociotopy in the temporoparietal cortex: common versus distinct processes. *Soc. Cognit. Affect. Neurosci.* 5, 48–58. <https://doi.org/10.1093/scan/nsp045>.
- Baker, B., Lansdell, B., Kording, K.P., 2022. Three aspects of representation in neuroscience. *Trends Cognit. Sci.* 26, 942–958. <https://doi.org/10.1016/j.tics.2022.08.014>.
- Bedny, M., Pascual-Leone, A., Saxe, R.R., 2009. Growing up blind does not change the neural bases of Theory of Mind. *Proc. Natl. Acad. Sci. U. S. A.* 106, 11312–11317. <https://doi.org/10.1073/pnas.0900010106>.
- Bellot, E., Abassi, E., Papeo, L., 2021. Moving toward versus away from another: how body motion direction changes the representation of bodies and actions in the visual cortex. *Cerebr. Cortex* 31, 2670–2685. <https://doi.org/10.1093/cercor/bhaa382>.
- Blake, R., Turner, L.M., Smoski, M.J., Pozdol, S.L., Stone, W.L., 2003. Visual recognition of biological motion is impaired in children with autism. *Psychol. Sci.* 14, 151–157. <https://doi.org/10.1111/1467-9280.01434>.

- Brass, M., Schmitt, R.M., Spengler, S., Gergely, G., 2007. Investigating action understanding: inferential processes versus action simulation. *Curr. Biol.* 17, 2117–2121. <https://doi.org/10.1016/j.cub.2007.11.057>.
- Brennan, J., Martin, E., 2012. Spatial proximity is more than just a distance measure. *Int. J. Hum. Comput. Stud.* 70, 88–106. <https://doi.org/10.1016/j.ijhcs.2011.08.006>.
- Brothers, L., Ring, B., Kling, A., 1990. Response of neurons in the macaque amygdala to complex social stimuli. *Behav. Brain Res.* 41, 199–213. [https://doi.org/10.1016/0166-4328\(90\)90108-Q](https://doi.org/10.1016/0166-4328(90)90108-Q).
- Buckholz, J.W., Meyer-Lindenberg, A., 2012. Psychopathology and the human connectome: toward a transdiagnostic model of risk for mental illness. *Neuron* 74, 990–1004. <https://doi.org/10.1016/j.neuron.2012.06.002>.
- Canessa, N., Alemanno, F., Riva, F., Zani, A., Proverbio, A.M., Mannara, N., Perani, D., Cappa, S.F., 2012. The neural bases of social intention understanding: the role of interaction goals. *PLoS One* 7, e42347. <https://doi.org/10.1371/journal.pone.0042347>.
- Caspers, S., Zilles, K., Laird, A.R., Eickhoff, S.B., 2010. ALE meta-analysis of action observation and imitation in the human brain. *Neuroimage* 50, 1148–1167. <https://doi.org/10.1016/j.neuroimage.2009.12.112>.
- Centelles, L., Assaiante, C., Nazarian, B., Anton, J.-L., Schmitz, C., 2011. Recruitment of both the mirror and the mentalizing networks when observing social interactions depicted by point-lights: a neuroimaging study. *PLoS One* 6, e15749. <https://doi.org/10.1371/journal.pone.0015749>.
- Chiavarino, C., Apperly, I.A., Humphreys, G.W., 2012. Understanding intentions: distinct processes for mirroring, representing, and conceptualizing. *Curr. Dir. Psychol. Sci.* 21, 284–289. <https://doi.org/10.1177/0963721412452727>.
- Ciaramidaro, A., Adenzato, M., Enrici, I., Erk, S., Pia, L., Bara, B.G., Walter, H., 2007. The intentional network: how the brain reads varieties of intentions. *Neuropsychologia* 45, 3105–3113. <https://doi.org/10.1016/j.neuropsychologia.2007.05>.
- de Borst, A.W., Sanchez-Vives, M.V., Slater, M., de Gelder, B., 2020. First-person virtual embodiment modulates the cortical network that encodes the bodily self and its surrounding space during the experience of domestic violence. *Eneuro* 7. <https://doi.org/10.1523/ENEURO.0263-19.2019>.
- de Gelder, B., Poyo Solanas, M., 2021. A computational neuroethology perspective on body and expression perception. *Trends Cognit. Sci.* 25, 744–756. <https://doi.org/10.1016/j.tics.2021.05.010>.
- de Gelder, B., Poyo Solanas, M., 2022. What postures communicate. *Enfance* 353–365. <https://doi.org/10.3917/enf2.223.0353>.
- de la Rosa, S., Choudhery, R.N., Curio, C., Ullman, S., Assif, L., Buelthoff, H.H., 2014. Visual categorization of social interactions. *Vis. Cognit.* 22, 1233–1271. <https://doi.org/10.1080/13506285.2014.991368>.
- Decety, J., Jackson, P.L., 2004. The functional architecture of human empathy. *Behav. Cognit. Neurosci. Rev.* 3, 71–100. <https://doi.org/10.1177/1534582304267187>.
- Deen, B., Koldewyn, K., Kanwisher, N., Saxe, R., 2015. Functional organization of social perception and cognition in the superior temporal sulcus. *Cerebr. Cortex* 25, 4596–4609. <https://doi.org/10.1093/cercor/bhv111>.
- Eisenberger, N.I., 2012. The neural bases of social pain: evidence for shared representations with physical pain. *Psychosom. Med.* 74, 126. <https://doi.org/10.1097/PSY.0b013e3182464dd1>.
- Feng, C., Eickhoff, S.B., Li, T., Wang, L., Becker, B., Camilleri, J.A., Hetu, S., Luo, Y., 2021. Common brain networks underlying human social interactions: evidence from large-scale neuroimaging meta-analysis. *Neurosci. Biobehav. Rev.* 126, 289–303.
- Fogassi, L., Ferrari, P.F., Gesierich, B., Rozzi, S., Chersi, F., Rizzolatti, G., 2005. Parietal lobe: from action organization to intention understanding. *Science* 308, 662–667. <https://doi.org/10.1126/science.1106138>.
- Frith, C.D., 2007. The social brain? *Philos. Trans. R. Soc. B Biol. Sci.* 362, 671–678. <https://doi.org/10.1098/rstb.2006.2003>.
- Frith, C.D., Frith, U., 2006. The neural basis of mentalizing. *Neuron* 50, 531–534. <https://doi.org/10.1016/j.neuron.2006.05.001>.
- Gallagher, H.L., Frith, C.D., 2003. Functional imaging of “theory of mind”. *Trends Cognit. Sci.* 7, 77–83. [https://doi.org/10.1016/S1364-6613\(02\)00025-6](https://doi.org/10.1016/S1364-6613(02)00025-6).
- Gallese, V., Fadiga, L., Fogassi, L., Rizzolatti, G., 1996. Action recognition in the premotor cortex. *Brain* 119, 593–609. <https://doi.org/10.1093/brain/119.2.593>.
- Gandolfo, M., Abassi, E., Balgova, E., Downing, P.E., Papeo, L., Koldewyn, K., 2024. Converging evidence that left extrastriate body area supports visual sensitivity to social interactions. *Curr. Biol.* 34, 343–351.e5. <https://doi.org/10.1016/j.cub.2023.12.009>.
- Giese, M.A., Poggio, T., 2003. Neural mechanisms for the recognition of biological movements. *Nat. Rev. Neurosci.* 4, 179–192. <https://doi.org/10.1038/nrn1057>.
- Gonzalez-Liencre, C., Zapata, L.E., Iruetagoiena, G., Seinfeld, S., Perez-Mendez, L., Arroyo-Palacios, J., Borland, D., Slater, M., Sanchez-Vives, M.V., 2020. Being the victim of intimate partner violence in virtual reality: first-versus third-person perspective. *Front. Psychol.* 11, 820. <https://doi.org/10.3389/fpsyg.2020.00820>.
- Grabenhorst, F., Rolls, E.T., 2011. Value, pleasure and choice in the ventral prefrontal cortex. *Trends Cognit. Sci.* 15, 56–67. <https://doi.org/10.1016/j.tics.2010.12.004>.
- Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., Blake, R., 2000. Brain areas involved in perception of biological motion. *J. Cognit. Neurosci.* 12, 711–720. <https://doi.org/10.1162/089929000562417>.
- Gu, R., Huang, W., Camilleri, J., Xu, P., Wei, P., Eickhoff, S.B., Feng, C., 2019. Love is analogous to money in human brain: coordinate-based and functional connectivity meta-analyses of social and monetary reward anticipation. *Neurosci. Biobehav. Rev.* 100, 108–128. <https://doi.org/10.1016/j.neubiorev.2019.02.017>.
- Hari, R., Kujala, M.V., 2009. Brain basis of human social interaction: from concepts to brain imaging. *Physiol. Rev.* 89, 453–479. <https://doi.org/10.1152/physrev.00041.2007>.
- Hasson, U., Nastase, S.A., Goldstein, A., 2020. Direct fit to nature: an evolutionary perspective on biological and artificial neural networks. *Neuron* 105, 416–434. <https://doi.org/10.1016/j.neuron.2019.12.002>.
- Heider, F., Simmel, M., 1944. An experimental study of apparent behavior. *Am. J. Psychol.* 57, 243–259. <https://doi.org/10.2307/1416950>.
- Hein, G., Knight, R.T., 2008. Superior temporal sulcus—it's my area: or is it? *J. Cognit. Neurosci.* 20, 2125–2136. <https://doi.org/10.1162/jocn.2008.20148>.
- 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, e111511. <https://doi.org/10.1371/journal.pone.0111511>.
- Iacoboni, M., Lieberman, M.D., Knowlton, B.J., Molnar-Szakacs, I., Moritz, M., Throop, C.J., Fiske, A.P., 2004. Watching social interactions produces dorsomedial prefrontal and medial parietal BOLD fMRI signal increases compared to a resting baseline. *Neuroimage* 21, 1167–1173. <https://doi.org/10.1016/j.neuroimage.2003.11.013>.
- 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 Biol.* 3, e79. <https://doi.org/10.1371/journal.pbio.0030079>.
- Isik, L., Koldewyn, K., Beeler, D., Kanwisher, N., 2017. Perceiving social interactions in the posterior superior temporal sulcus. *Proc. Natl. Acad. Sci. U. S. A.* 114, E9145–E9152. <https://doi.org/10.1073/pnas.1714471114>.
- Izuma, K., Saito, D.N., Sadato, N., 2008. Processing of social and monetary rewards in the human striatum. *Neuron* 58, 284–294. <https://doi.org/10.1016/j.neuron.2008.03.020>.
- Jacob, P., Jeannerod, M., 2005. The motor theory of social cognition: a critique. *Trends Cognit. Sci.* 9, 21–25. <https://doi.org/10.1016/j.tics.2004.11.003>.
- Jellema, T., Baker, C., Wicker, B., Perrett, D., 2000. Neural representation for the perception of the intentionality of actions. *Brain Cognit.* 44, 280–302. <https://doi.org/10.1006/brcg.2000.1231>.
- Keck, J., Zabicki, A., Bachmann, J., Munzert, J., Krüger, B., 2022. Decoding spatiotemporal features of emotional body language in social interactions. *Sci. Rep.* 12, 15088. <https://doi.org/10.1038/s41598-022-19267-5>.
- Kiley Hamlin, J., Ullman, T., Tenenbaum, J., Goodman, N., Baker, C., 2013. The mentalistic basis of core social cognition: experiments in preverbal infants and a computational model. *Dev. Sci.* 16, 209–226. <https://doi.org/10.1111/desc.12017>.
- Kokkinara, E., Slater, M., López-Moliner, J., 2015. The effects of visuomotor calibration to the perceived space and body, through embodiment in immersive virtual reality. *Trans. Appl. Percept.* 13, 1–22. <https://doi.org/10.1145/2818998>.
- Konrad, A.C., Förster, K., Stretton, J., Dalgleish, T., Böckler-Raettig, A., Trautwein, F.M., Singer, T., Kanske, P., 2024. Risk factors for internalizing symptoms: the influence of empathy, theory of mind, and negative thinking processes. *Hum. Brain Mapp.* 45, e26576. <https://doi.org/10.1002/hbm.26576>.

- Kujala, M.V., Carlson, S., Hari, R., 2012. Engagement of amygdala in third-person view of face-to-face interaction. *Hum. Brain Mapp.* 33, 1753–1762. <https://doi.org/10.1002/hbm.21317>.
- Landsiedel, J., Daughters, K., Downing, P.E., Koldewyn, K., 2022. The role of motion in the neural representation of social interactions in the posterior temporal cortex. *Neuroimage* 262, 119533. <https://doi.org/10.1016/j.neuroimage.2022.119533>.
- Lavelle, M., Healey, P.G., McCabe, R., 2014. Nonverbal behavior during face-to-face social interaction in schizophrenia: a review. *J. Nerv. Ment. Dis.* 202, 47–54. <https://doi.org/10.1097/NMD.000000000000031>.
- Levy, D.J., Glimcher, P.W., 2012. The root of all value: a neural common currency for choice. *Curr. Opin. Neurobiol.* 22, 1027–1038. <https://doi.org/10.1016/j.conb.2012.06.001>.
- Li, B., Poyo Solanas, M., Marrazzo, G., Raman, R., Taubert, N., Giese, M., Vogels, R., de Gelder, B., 2023. A large-scale brain network of species-specific dynamic human body perception. *Prog. Neurobiol.* 221, 102398. <https://doi.org/10.1016/j.pneurobio.2022.102398>.
- Li, B., Poyo Solanas, M., Marrazzo, G., de Gelder, B., 2024. Connectivity and functional diversity of different temporo-occipital nodes for action perception. *bioRxiv*. <https://doi.org/10.1101/2024.01.12.574860>, 2024.2001. 2012.574860.
- Lin, A., Adolphs, R., Rangel, A., 2012. Social and monetary reward learning engage overlapping neural substrates. *Soc. Cognit. Affect. Neurosci.* 7, 274–281. <https://doi.org/10.1093/scan/nsr006>.
- Malik, M., Isik, L., 2023. Relational visual representations underlie human social interaction recognition. *Nat. Commun.* 14, 7317. <https://doi.org/10.1038/s41467-023-43156-8>.
- Manera, V., Del Giudice, M., Bara, B.G., Verfaillie, K., Becchio, C., 2011. The second-agent effect: communicative gestures increase the likelihood of perceiving a second agent. *PLoS One* 6, e22650. <https://doi.org/10.1371/journal.pone.0022650>.
- Marrazzo, G., Vaessen, M.J., de Gelder, B., 2021. Decoding the difference between explicit and implicit body expression representation in high level visual, prefrontal and inferior parietal cortex. *Neuroimage* 243, 118545. <https://doi.org/10.1016/j.neuroimage.2021.118545>.
- Masson, H.L., Isik, L., 2021. Functional selectivity for social interaction perception in the human superior temporal sulcus during natural viewing. *Neuroimage* 245, 118741. <https://doi.org/10.1016/j.neuroimage.2021.118741>.
- McMahon, E., Isik, L., 2023. Seeing social interactions. *Trends Cognit. Sci.* <https://doi.org/10.1016/j.tics.2023.09.001>.
- McMahon, E., Bonner, M.F., Isik, L., 2023. Hierarchical organization of social action features along the lateral visual pathway. *Curr. Biol.* 33, 5035–5047. <https://doi.org/10.1016/j.cub.2023.10.015>.
- Mello, M., Dupont, L., Engelen, T., Acciarino, A., de Borst, A.W., de Gelder, B., 2022. The influence of body expression, group affiliation and threat proximity on interactions in virtual reality. *Curr. Res. Behav. Sci.* 3, 100075. <https://doi.org/10.1016/j.crbeha.2022.100075>.
- Miller, E.K., Cohen, J.D., 2001. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202. <https://doi.org/10.1146/annurev.neuro.24.1.167>.
- Molenberghs, P., Cunnington, R., Mattingley, J.B., 2012. Brain regions with mirror properties: a meta-analysis of 125 human fMRI studies. *Neurosci. Biobehav. Rev.* 36, 341–349. <https://doi.org/10.1016/j.neubiorev.2011.07.004>.
- Neri, P., Luu, J.Y., Levi, D.M., 2006. Meaningful interactions can enhance visual discrimination of human agents. *Nat. Neurosci.* 9, 1186–1192. <https://doi.org/10.1038/nn1759>.
- Papeo, L., 2020. Twos in human visual perception. *Cortex* 132, 473–478. <https://doi.org/10.1016/j.cortex.2020.06.005>.
- Papeo, L., Goupil, N., Soto-Faraco, S., 2019. Visual search for people among people. *Psychol. Sci.* 30, 1483–1496. <https://doi.org/10.1177/0956797619867295>.
- Paxton, A., Dale, R., 2017. Interpersonal movement synchrony responds to high-and low-level conversational constraints. *Front. Psychol.* 8, 1135. <https://doi.org/10.3389/fpsyg.2017.01135>.
- Pitcher, D., Ungerleider, L.G., 2021. Evidence for a third visual pathway specialized for social perception. *Trends Cognit. Sci.* 25, 100–110. <https://doi.org/10.1016/j.tics.2020.11.006>.
- Poyo Solanas, M., Vaessen, M., de Gelder, B., 2020a. Computation-based feature representation of body expressions in the human brain. *Cerebr. Cortex* 30, 6376–6390. <https://doi.org/10.1093/cercor/bhaa196>.
- Poyo Solanas, M., Vaessen, M.J., de Gelder, B., 2020b. The role of computational and subjective features in emotional body expressions. *Sci. Rep.* 10, 6202. <https://doi.org/10.1038/s41598-020-63125-1>.
- Proverbio, A.M., Riva, F., Paganelli, L., Cappa, S.F., Canessa, N., Perani, D., Zani, A., 2011. Neural coding of cooperative vs. affective human interactions: 150 ms to code the action's purpose. *PLoS One* 6, e22026. <https://doi.org/10.1371/journal.pone.0022026>.
- Rangel, A., Camerer, C., Montague, P.R., 2008. A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* 9, 545–556. <https://doi.org/10.1038/nrn2357>.
- Rizzolatti, G., Craighero, L., 2004. The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192. <https://doi.org/10.1146/annurev.neuro.27.070203.144230>.
- Rizzolatti, G., Sinigaglia, C., 2010. The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nat. Rev. Neurosci.* 11, 264–274. <https://doi.org/10.1038/nrn2805>.
- Rizzolatti, G., Fogassi, L., Gallese, V., 2001. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci.* 2, 661–670. <https://doi.org/10.1038/35090060>.
- Sanchez-Vives, M.V., Slater, M., 2005. From presence to consciousness through virtual reality. *Nat. Rev. Neurosci.* 6, 332–339. <https://doi.org/10.1038/nrn1651>.
- Sapey-Triomphe, L.-A., Centelles, L., Roth, M., Fonlupt, P., Henaff, M.-A., Schmitz, C., Assaiante, C., 2017. Deciphering human motion to discriminate social interactions: a developmental neuroimaging study. *Soc. Cognit. Affect. Neurosci.* 12, 340–351. <https://doi.org/10.1093/scan/nsw117>.
- Saxe, R., 2005. Against simulation: the argument from error. *Trends Cognit. Sci.* 9, 174–179. <https://doi.org/10.1016/j.tics.2005.01.012>.
- Saxe, R., 2006. Uniquely human social cognition. *Curr. Opin. Neurobiol.* 16, 235–239. <https://doi.org/10.1016/j.conb.2006.03.001>.
- Saxe, R., Haushofer, J., 2008. For love or money: a common neural currency for social and monetary reward. *Neuron* 58, 164–165. <https://doi.org/10.1016/j.neuron.2008.04.005>.
- Schultz, J., Friston, K.J., O'Doherty, J., Wolpert, D.M., Frith, C.D., 2005. Activation in posterior superior temporal sulcus parallels parameter inducing the percept of animacy. *Neuron* 45, 625–635. <https://doi.org/10.1016/j.neuron.2004.12.052>.
- Seinfeld, S., Zhan, M., Poyo-Solanas, M., Barsuola, G., Vaessen, M., Slater, M., Sanchez-Vives, M.V., de Gelder, B., 2021. Being the victim of virtual abuse changes default mode network responses to emotional expressions. *Cortex* 135, 268–284. <https://doi.org/10.1016/j.cortex.2020.11.018>.
- Shu, T., Bhandwaldar, A., Gan, C., Smith, K., Liu, S., Gutfreund, D., Spelke, E., Tenenbaum, J., Ullman, T., 2021. Agent: a benchmark for core psychological reasoning. In: Paper presented at: International Conference on Machine Learning (PMLR).
- Sinke, C.B., Sorger, B., Goebel, R., de Gelder, B., 2010. Tease or threat? Judging social interactions from bodily expressions. *Neuroimage* 49, 1717–1727. <https://doi.org/10.1016/j.neuroimage.2009.09.065>.
- Slater, M., 2009. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 3549–3557. <https://doi.org/10.1098/rstb.2009.0138>.
- Smallwood, J., Bernhardt, B.C., Leech, R., Bzdok, D., Jefferies, E., Margulies, D.S., 2021. The default mode network in cognition: a topographical perspective. *Nat. Rev. Neurosci.* 22, 503–513. <https://doi.org/10.1038/s41583-021-00474-4>.
- Spunt, R.P., Lieberman, M.D., 2012a. Dissociating modality-specific and supramodal neural systems for action understanding. *J. Neurosci.* 32, 3575–3583. <https://doi.org/10.1523/JNEUROSCI.5715-11.2012>.
- Spunt, R.P., Lieberman, M.D., 2012b. An integrative model of the neural systems supporting the comprehension of observed emotional behavior. *Neuroimage* 59, 3050–3059. <https://doi.org/10.1016/j.neuroimage.2011.10.005>.
- Spunt, R.P., Kemmerer, D., Adolphs, R., 2016. The neural basis of conceptualizing the same action at different levels of abstraction. *Soc. Cognit. Affect. Neurosci.* 11, 1141–1151. <https://doi.org/10.1093/scan/nsv084>.

- Su, J., Van Boxtel, J.J., Lu, H., 2016. Social interactions receive priority to conscious perception. *PLoS One* 11, e0160468. <https://doi.org/10.1371/journal.pone.0160468>.
- Tager-Flusberg, H., 2007. Evaluating the theory-of-mind hypothesis of autism. *Curr. Dir. Psychol. Sci.* 16, 311–315. <https://doi.org/10.1111/j.1467-8721.2007.00527.x>.
- Tarhan, L., Konkle, T., 2020. Sociality and interaction envelope organize visual action representations. *Nat. Commun.* 11, 3002. <https://doi.org/10.1038/s41467-020-16846-w>.
- Ullman, T., Baker, C., Macindoe, O., Evans, O., Goodman, N., Tenenbaum, J., 2009. Help or hinder: Bayesian models of social goal inference. *Adv. Neural Inf. Process. Syst.* 22.
- Van den Stock, J., Hortensius, R., Sinke, C., Goebel, R., de Gelder, B., 2015. Personality traits predict brain activation and connectivity when witnessing a violent conflict. *Sci. Rep.* 5, 13779. <https://doi.org/10.1038/srep13779>.
- Vestner, T., Tipper, S.P., Hartley, T., Over, H., Rueschemeyer, S.-A., 2019. Bound together: social binding leads to faster processing, spatial distortion, and enhanced memory of interacting partners. *J. Exp. Psychol. Gen.* 148, 1251. <https://doi.org/10.1037/xge0000545>.
- Walbrin, J., Koldewyn, K., 2019. Dyadic interaction processing in the posterior temporal cortex. *Neuroimage* 198, 296–302. <https://doi.org/10.1016/j.neuroimage.2019.05.027>.
- Walbrin, J., Downing, P., Koldewyn, K., 2018. Neural responses to visually observed social interactions. *Neuropsychologia* 112, 31–39. <https://doi.org/10.1016/j.neuropsychologia.2018.02.023>.
- Weiner, K.S., Grill-Spector, K., 2011. Not one extrastriate body area: using anatomical landmarks, hMT+, and visual field maps to parcellate limb-selective activations in human lateral occipitotemporal cortex. *Neuroimage* 56, 2183–2199. <https://doi.org/10.1016/j.neuroimage.2011.03.041>.
- 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, 655–664. [https://doi.org/10.1016/S0896-6273\(03\)00679-2](https://doi.org/10.1016/S0896-6273(03)00679-2).
- Wurm, M.F., Caramazza, A., 2022. Two “what” pathways for action and object recognition. *Trends Cognit. Sci.* 26, 103–116. <https://doi.org/10.1016/j.tics.2021.10.003>.