The neuroscience of people watching

Quadflieg, Susanne; Koldewyn, Kami

Annals of the New York Academy of Sciences

09. Oct. 2020

The Neuroscience of People-Watching:

How the Human Brain Makes Sense of Other People's Encounters

Susanne Quadflieg[1] & Kami Koldewyn[2]

[1]School of Experimental Psychology, University of Bristol, Bristol, UK

[2]School of Psychology, Bangor University, Bangor, UK

ABSTRACT

Neuroscientific investigations interested in questions of person perception and impression formation traditionally asked their participants to observe and evaluate isolated individuals. In recent years, however, there has been a surge of studies presenting third-party encounters between two (or more) individuals as stimuli. Due to this subtle methodological change, the brain's capacity to understand other people's interactions and relationships from limited visual information – also known as people-watching – has become a distinct topic of inquiry. Though initial evidence indicates this capacity relies on several well-known networks of the social brain (including the person perception network, the action observation network, and the mentalizing network), a comprehensive framework of people-watching must overcome three major challenges. First, it must develop a taxonomy of judgments people habitually make when witnessing the encounters of others. Second, it must clarify which visual cues give rise to these encounter-based judgments. Third, it must elucidate how and why several brain networks work together to accomplish these judgments. To advance all three lines of research, the current article summarizes what is currently known, but also what remains to be studied about the neuroscience of people-watching.

Keywords: person perception; social cognition; social interaction; social neuroscience; third-person perspective

According to the social intelligence hypothesis, the ability to quickly detect and understand intricate social relations between other people may have facilitated the evolutionary development of unusually large brains in humans.[1] Despite this fascinating claim, cognitive neuroscientists have rarely studied the brain's response to relations between people. Instead, much of their work has focused on the perception and evaluation of isolated individuals or their parts, such as a human face or body. Without doubt, this approach has resulted in numerous seminal insights. It has revealed, for instance, that brain regions such as the fusiform face area (FFA) or the extrastriate body area (EBA) are specifically tuned for encoding the visual appearance of human faces and bodies.[2,3] But it has not examined how the human brain analyzes scenarios involving multiple individuals, including their social interactions and relationships.

This lack of systematic inquiry may seem surprising, considering that making sense of other people's encounters signifies a fascinating human capacity.[4] As any avid people-watcher can attest, observing strangers in each other's company rapidly prompts numerous social judgments about how and why they have come together.[5,6] In acknowledgement of this phenomenon, psychologists have long studied social impressions from so-called third-party encounters (TPEs). By contrast, neuroscientific research on the perception and interpretation of TPEs has largely been absent. During the last five years, however, photographs and video clips of social scenarios including two or more individuals (see Figure 1) have begun to feature regularly in studies using event-related potential[7,8] (ERP), functional magnetic resonance imaging[9,10] (fMRI), and transcranial magnetic stimulation[11,12] (TMS). But which innovative insights, if any, has this new approach uncovered?

To address this question, the current article reflects on a growing body of work that explores the neural substrates of forming impressions from other people's encounters. In so doing, the article examines the human brain's ability to analyze and interpret the appearance and actions of multiple individuals simultaneously in order to understand the social interactions and relation between them. The reader is first introduced to encounter-based impressions from a psychological perspective. Subsequently, experimental studies that have traced neural responses to TPEs in various well-known brain networks are discussed. Finally, initial attempts to understand neural variation in response to TPEs in clinical populations will be summarized. By providing an overview of recent theoretical and empirical approaches on the neural substrates of encounter-based impressions, this article aims to highlight their influential role in human social cognition.
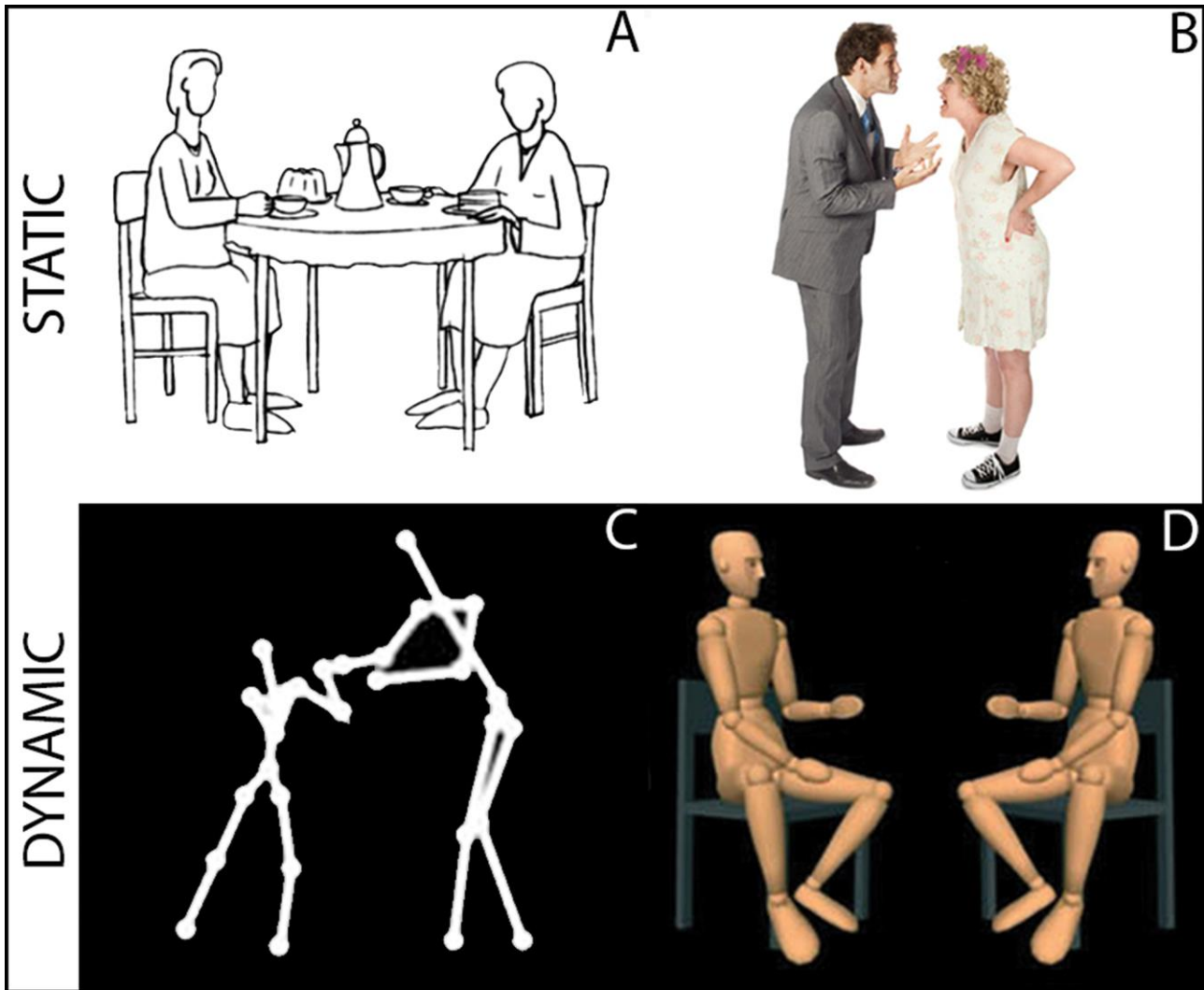
Figure 1. Portrayals of person dyads as used in recent neuroscientific studies, ranging from static A) drawings (Source: ref 67. Copyright 2010 Elsevier) and B) photographs (Source: ref 10. Copyright 2015 Elsevier) to dynamic C) point-light displays (Source: ref 9. Copyright 2014 John Wiley & Sons Ltd) and D) stick figures (Source: ref 71. Copyright 2014 John Wiley & Sons Ltd). All images reprinted with permission.

In close adherence to its agenda, the article does not cover the formation of encounter-based impressions in response to hearsay (i.e., in response to verbal descriptions of human encounters received from other people[13,14]), nor the formation of anthropomorphic impressions in response to non-human entities (such as geometric shapes) purposefully made to 'interact' in a human-like manner.[15-18] Although both lines of research probe important variants of encounter-based impressions, they necessarily skip the role played by basic perceptual operations dedicated towards face and body encoding and how these operations determine the course and products of the impression formation process – a question that lies very much at the heart of the phenomenon known as people-watching.[19,20]

People-Watching: A Psychological Perspective

Over fifty years ago, psychologists began to study how the human mind makes sense of encounters between multiple individuals from a third-person perspective.[5,21,22] This early work focused primarily on the perception and evaluation of person dyads as it had been noticed that humans gather primarily in sets of two in public places[23] and spend about half of their social time in the company of just one other person.[24,25] Besides acknowledging the importance of person dyads in human life, early research on encounter-based impressions was fueled by the realization that these impressions could directly affect observers' own behavior. It was demonstrated, for instance, that in busy streets most individuals refrain from penetrating the space between two people whom they consider a meaningful social unit.[26,27] What was less clear at the time, and what remains a topic of contemporary debate, is the question of what exactly constitutes such a unit from an observer's point of view?

According to numerous behavioral studies, meaningful social units are usually detected on the basis of overt markers of interpersonal involvement between people, ranging from shared eye gaze, posture mimicry, close physical proximity, and movement synchrony to direct communicative gestures and speech acts.[28,29] Yet this prevalent approach of studying the observation and interpretation of human encounters has prioritized the assessment of momentary social interactions over other types of social relations. As a case in point, imagine two people who are waiting at the same bus stop, but who are not currently interacting with, or attending towards, one another. Despite this lack of direct involvement, witnessing both individuals together can elicit unique relational social insights. Both individuals may, for instance, look like they share a family resemblance, have a similar racial background, and/or support the same sports team.[30] The similarity of their actions (e.g., waiting) may further reveal that they share a common goal or destiny (i.e., catching the bus). In short, observing people in each other's company can prompt numerous inferences that go beyond mere judgments of direct interpersonal involvement.[31]

In acknowledgement of this circumstance, contemporary psychologists describe and investigate encounter-based impressions as a multifaceted phenomenon that entails a wide range of *perceptual, action,* and *social* appraisals.[4] In terms of perceptual appraisals, for instance, observers of TPEs seem to quickly assess whether co-occurring individuals display salient visual markers of social relatedness, such as mutual smiles, coordinated movement patterns, or shared social group memberships.[29,32,33] In terms of action appraisals, they regularly examine whether TPEs entail individuals who engage in independent or joint

actions[34] (e.g., reading vs. chatting), in goal-compatible or incompatible actions[35] (e.g., collaborating vs. competing), and/or in positive or negative actions[36] (e.g., kissing vs. pushing someone). In terms of social appraisals, finally, observers often speculate about a TPE's momentary level of formality, intimacy, rapport, and subordination/domination[5,37,38] as well as about people's overarching type of acquaintance[39,40] (e.g., whether co-occurring individuals are strangers, colleagues, friends etc.).

Although it remains uncertain why TPEs prompt such a wide range of inferences in uninvolved bystanders, it has previously been argued that the human inclination to analyze TPEs in elaborate detail may have profited from evolutionary pressures.[41,42] Considering that humans must interact with both kin and non-kin individuals to survive in the face of adversity[43], the careful analysis of TPEs may have helped individuals to decide whom to choose as their own interaction partners[44] (e.g., someone with a history of mutually beneficial social exchanges). Additionally or alternatively, TPEs may have offered (and remain to offer) a unique opportunity for observational learning, allowing those who witness them to widen their own behavioral repertoire without taking personal risks while doing so.[45-48] Yet regardless of which factors may ultimately have facilitated the emergence of encounter-based impressions in humans, their unique scope and nature makes these impressions undoubtedly a prevalent aspect of social cognition in present-day life. But what do we understand at this point about how the human brain implements these impressions? Have recent neuroscientific insights on encounter-based impressions helped to falsify or refine existing psychological theories about them?[49] Or have such insights inspired a novel framework on how the human brain encodes and analyses visual information about other people's encounters?

People-Watching: Towards A Neuroscientific Perspective

According to traditional neuroscientific investigations, three brain networks play a pivotal role when people observe and evaluate each other:[19,50] the person perception network (PPN), the action observation network (AON), and the mentalizing network (MTN). When watching isolated individuals, the PPN is believed to accomplish the visual analysis of other people's faces and bodies[51,52], the AON is considered to decipher and predict other people's actions[53,54], and the MTN is understood to infer other people's invisible mental states (including their beliefs, desires, motives, or attitudes) and personality traits.[55,56] Although numerous studies indicate that all three networks also play a pivotal role during the

observation and interpretation of TPEs[57-60], their functional significance in the context of encounter-based impressions is less well understood.

This lack of understanding is largely due to the fact that existing neuroscientific studies on TPE processing vary substantially in their methods. In terms of stimuli, for instance, researchers have used various types of media to present TPEs over the years, ranging from still images[60,61] to brief sequences of still images[62,63] to dynamic video clips.[57,64] Furthermore, still images have included color photographs[6,10], grayscale photographs[65,66], and black-and-white schematic drawings.[67,68] Video clips, in addition, have comprised realistic depictions of social interactions[69,70], digital animations of human-like avatars[59,62], and point-light displays as well as stick-figure displays of human movements.[9,71] Finally, whereas many stimuli portrayed target individuals from head/neck to toe[10,12], others showed only people's upper bodies[57,66], facial expressions[72], or hand movements.[11,73]

Each of these different sets of stimuli comes with its own limitation(s). Generally speaking, when forming impressions from static full body photographs (see Figure 1B), observers can learn whether two (or more) target individuals look alike, are in close physical proximity, mimic each other's expressions and postures, and engage in shared eye contact, interpersonal touch, or direct communication via gestures or speech. But only from dynamic portrayals can they extract the frequency, duration, and coordination of various nonverbal events (e.g., reciprocated smiles) and the degree of motion synchrony and turn-taking between people. Given the diversity of stimuli used, it may not come as a surprise to learn that there has been little overlap in terms of brain activity during TPE exposure across studies. This lack of overlap suggests, however, that the presence (or absence) of certain visual markers fundamentally affects the impression formation process and its corresponding neural signature. In further support of this claim, it has also been shown that the exact same social interactions can prompt rather different patterns of brain activity depending on whether the agents' eye gaze is visible or not.[66]. In light of these findings, both neuroscientific as well as psychological theories on encounter-based impressions should strive to define more clearly how various face, body, and motion cues that commonly characterize TPEs can affect the course and outcome of perceivers' inferences.[32]

Aside from using various types of stimuli, existing neuroscientific studies on TPE processing have also differed substantially in terms of their experimental conditions of interest, including their so-called "baseline" condition. Whereas some scientists have compared the neural effects of TPEs with those elicited by non-social control displays[61,74,75],

others have contrasted TPEs and depictions of single individuals[11,12,76,77] or TPEs of jointly acting individuals with TPEs of independently acting individuals.[58,69,78] Beyond these three major lines of research, various subtypes of joint actions (i.e., interactions) have been pitched against each other. Neural responses have been measured, for example, for interactions characterized by contingent or non-contingent movements between interaction partners[71], for interactions of positive or negative valence[70,79], and for interactions serving primarily instrumental or affiliative goals.[6,7] Again, this varied approach has somewhat hindered accumulative insights on how the brain creates encounter-based impressions. Yet by trying to understand the neural effects of different kinds of TPEs, neuroscientists have begun to examine common properties of human encounters (e.g., their degree of instrumentality) that have gone largely unstudied by psychologists. As such, their work indicates that an inclusive theory on TPE processing must integrate both psychological and neuroscientific lines of research to derive a truly comprehensive taxonomy of encounter-based impressions.

In this context, it is also worth noting that the importance of perceivers' processing goals during TPE observation has attracted more attention in recent neuroscientific studies than in traditional psychological studies. Specifically, neuroscientific studies have prompted their participants to judge TPEs along various different dimensions of relevance, such as their perceptual attributes (e.g., Do both individuals have the same gender?[10] Which individual looks heavier than the other?[65]), their action properties (e.g., Do both individuals act jointly or independently?[58,66]), or their social qualities (e.g., Is one individual threatening the other?[80] Is one individual helping the other?[81]). They have further asked participants to simply view TPEs without specific instructions[57,60,64,78] or to complete TPE-unrelated tasks (e.g., using a standard one-back detection task[9] or a counting paradigm[82]). Most importantly, there have even been initial attempts to directly compare the effects of different impression formation tasks in the same study. In consequence, it has been shown that explicit social judgments relative to basic perceptual tasks or TPE-unrelated tasks produce systematically enhanced activity in the PPN, AON, as well as the MTN.[65,68,80] Additionally, judgments about *how* an encounter unfolds (e.g., about which actions it entails) seem to result in stronger AON activation than judgments about *why* it unfolds (e.g., what types of motives or intentions people have). Vice versa, why-related judgments appear to recruit the MTN more strongly than how-related judgments.[79,83] In combination, these findings suggest that not only

variations in people's actual observations, but also in the goals that guide these observations shape the impression formation process.

In summary, neuroscientific research on TPEs has progressed in many directions during the last 5 years. This progress has raised several important issues overlooked by prior psychological theorizing, but has not yet inspired a coherent neuroscientific framework instead. In particular, the diversity of stimuli used and impressions probed by contemporary neuroscientists make it hard to determine to which extent the latest neuroscientific findings can be compared with one another and/or generalize to people-watching in the real world. In consequence, further work is required to understand how people-watchers integrate a wide range of visual cues during naturalistic observations of human encounters to form far-reaching impressions about them. To encourage more systematic research in the field, the effects of people-watching on neural activity in the PPN, AON, and MTN are discussed in further detail below. By looking at the neural effects of TPE processing in well-established brain networks, we aim to discover initial patterns of converging evidence that can form a promising starting point for future investigations.

TPE Processing in the Person Perception Network

Decades of neuroimaging research have revealed that simply looking at another person recruits a specific set of brain regions widely known as the PPN.[84,85] The PPN generally allows humans to detect, encode, and analyze the complex visual information that constitutes a person's idiosyncratic appearance and way of movement. The network's key nodes (see Figure 2A) are primarily found in the brain's ventral visual processing stream and respond selectively to human faces[86], human bodies[87], and human motion.[88] Depending on which aspect of a person is visible at any given time (i.e., a person's face and/or body at rest or in motion), activity increases in the PPN may be observed in the occipital face area (OFA), fusiform face area (FFA), fusiform body area (FBA), extrastriate body area (EBA), and/or the posterior superior temporal sulcus (pSTS; a region sensitive to dynamic facial and bodily input).

Despite abundant neuroimaging work on how the PPN extracts and recognizes a person's idiosyncratic appearance[89,90], very little work has explored the network's contribution to processing encounters between multiple people. Indeed, in comparison to the other two networks discussed in this review, the PPN may have received the least scrutiny in the context of TPEs. This is hardly surprising, considering that researchers interested in this

network have just begun to explore how it integrates visual input extracted from different body parts (e.g., how it combines facial and bodily information[91,92]). Under these circumstances, aiming to understand how visual information is integrated across multiple individuals (if at all) makes for a daunting prospect. At the same time, however, behavioral findings strongly suggest that detecting the presence of a person is facilitated when s/he is engaged in a meaningful interaction with another.[29,34] These data indicate that even very basic aspects of person perception may be altered when multiple individuals are observed simultaneously.
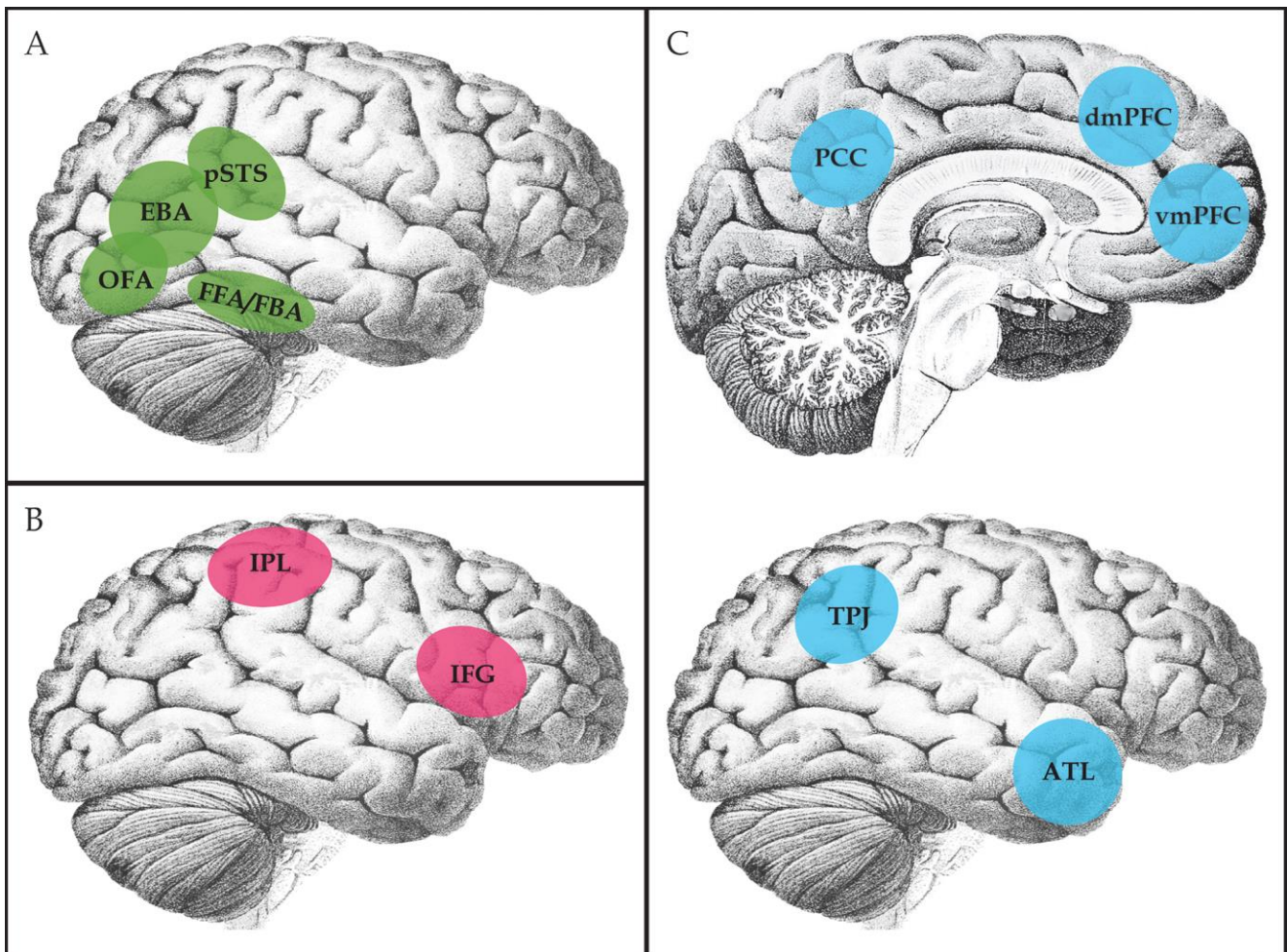
Figure 2. A schematic depiction of brain regions forming the core person perception network (A), the action observation network (B), and the mentalizing network (C) as discussed in this article.

This possibility raises a number of important questions. For instance, do the same brain regions that implement the visual analysis of isolated individuals also contribute to the perception of encounters between individuals? If so, to which degree (if at all) is their activity

affected by pivotal perceptual, action, and/or social attributes of the encounter itself (e.g., its degree of movement coordination, goal compatibility, and/or level of rapport?). Finally, do regions of the PPN process information from distinct individuals separately and require 'higher-order' networks to combine this information or does the PPN itself show signs of integration of visual information across individuals? Although most of these questions remain unanswered to date, their relevance becomes apparent in light of findings showing that observing TPEs reliably activates the PPN.[10,58-60,64,67,77,80,93]

Yet very few of the studies which found PPN activity in response to TPEs in the past were explicitly designed to measure it. As a result, several of them compared the processing of scenes with multiple individuals to the processing of scenes without people at all, making it difficult to determine whether the obtained results reflect the observation of human encounters or of people per se[93]. In addition, of those studies that explicitly contrasted portrayals of TPEs with portrayals of isolated individuals, none actually identified regions of the PPN with an independent localizer task.[64,67] As such, it remains somewhat speculative whether the activations reported (e.g., in the fusiform gyrus or the superior temporal sulcus) are <u>actually</u> part of the traditional PPN or not. Despite these challenges, converging evidence begins to indicate that the observation of TPEs compared to the observation of isolated individuals enhances activity in two well-defined regions of the PPN, namely the EBA[59,73,78,] and the pSTS.[59,64,67] Further evidence suggests that observing multiple individuals who engage in joint actions (compared to individuals who engage in independent actions) also elicits an enhanced response in these two regions.[57,58] <u>In comparison</u>, these types of contrasts have rarely resulted in modulations of activity in other regions of the PPN, such as the OFA, FFA, or FBA (but see ref 58 for fusiform activity towards joint actions).

For many neuroscientists, the prominent role of the pSTS during the observation of TPEs does not come as a surprise. The region is often described as part of the PPN, as well as the AON and MTN, and has been declared a "hub" of the social brain that implements not only the visual analysis of conspecifics, but also contributes to interpreting their actions and internal mental states.[50,90,94] Further support for the region's involvement in TPE processing comes from one very recent study that looked at brain activity in response to a 30 minute movie excerpt. This study adopted a data-driven "reverse correlation" analysis approach and found that, along with other regions, cortex in and around both the left and the right pSTS showed a preference for portrayals of multiple individuals over single-person scenes.[95] Importantly, these findings closely resemble earlier results obtained with static line

drawings[67] and artificial avatars.[59] Together, they support the notion that the pSTS directly contributes to the analysis of scenarios that comprise multiple individuals.

What is less certain is whether activity in the pSTS is mainly sensitive to the number of people visible or to the actual interactions and relations between them. At least one study that localized the pSTS, for instance, did not manage to observe differential activity in this region for still pictures of person dyads that showed people either facing towards or away from each[10] (but see ref 60 for counter-evidence). Yet there is initial evidence that the pSTS distinguishes between dynamic portrayals of person dyads that comprise two people who engage in joint versus independent actions.[57,58] Further work has revealed that interactions characterized by contingent rather than non-contingent actions elicit stronger pSTS activity[71] and that pSTS responses may be increased in response to negative rather than neutral interactions.[63,67] According to additional data, however, comparing negative and positive encounters does not necessarily result in differential pSTS activity[61,79], suggesting that the region may respond to valenced actions per se rather than negativity in particular. To better understand the functional significance of the pSTS in future, however, the type of variations in TPEs that can modulate its activity require further investigation. Such work should also begin to contrast the effects of dynamic and static TPE portrayals. Given that the pSTS is generally much more engaged by dynamic than static stimuli[90,96], its full sensitivity towards different types of TPEs may only become apparent when videos rather than still images are used for investigation.

Equally deserving of further examination are the conditions under which the EBA responds to human encounters. At least three studies have suggested increases in this region whenever two people engage in actions that are incompatible with one another.[9,10,80] For example, bilateral increases in the EBA have been reported in response to scenarios that involve one individual intending to take hold of another who, in turn, is trying to take flight.[80] A similar data pattern has been reported in response to interactions that involve mismatching actions between people[10] (e.g., one person trying to high-five another who intends to shake hands). These findings tentatively suggest that the EBA may generate perceptual predictions about compatible body postures and movements between people that result in enhanced processing when these predictions are violated. Interestingly, such an interpretation challenges the notion that the EBA simply encodes the configuration of human bodies in order to enable other cortical systems to interpret this information in its social

context[97] and argues instead that the region closely interacts with other systems to structure incoming percepts based on prior experience and stored social knowledge.[98,99]

In conclusion, although initial findings suggest that at least some nodes of the PPN are vitally involved in processing TPEs, work investigating this phenomenon is still in its infancy. To draw stronger conclusions about whether and how specific nodes in the PPN are involved in the analysis of other people's encounters, future work should aim to identify these regions in each participant's brain using standardized localizer tasks before testing their response to various kinds of TPEs. These TPEs should be carefully controlled for low-level visual differences across experimental conditions (e.g., the visibility of facial expressions, full-body shapes, and the level of implied or actually portrayed motion) and participants should be prompted to process each encounter with a standardized processing goal in mind. Finally, researchers should begin to investigate whether the PPN response to various kinds of human encounters can be predicted by an unweighted or weighted sum of the response to the constituent individuals and/or whether it shows non-additive properties.[100,101]

TPE Processing in the Action Observation Network (AON)

When observing a person in action (i.e., during the exhibition of intentional motor behavior) rather than at rest, not only the PPN gets recruited, but also the AON.[102-105] The AON is widely considered a brain network dedicated towards action understanding and its key nodes include the inferior parietal lobule (IPL), the inferior frontal gyrus (IFG), and the ventral premotor cortex adjacent to the IFG[106-108] (Figure 2B). Though the network's role during TPE processing also requires further examination, there is initial evidence that dynamic portrayals of TPEs compared to equivalent portrayals of isolated individuals enhance activity in the right IFG and the right premotor cortex.[59,64,78] As such, there is reason to believe that the AON distinguishes systematically between human actions that involve one or multiple actors.

There is further evidence that watching TPEs of multiple individuals who engage in joint actions recruits the IFG more strongly than observing TPEs of independently acting individuals[57,58,66] (see Figure 3). These data support the assumption that the IFG is particularly involved in understanding the coordination of actions between individuals. Additional findings substantiating this claim have revealed that the IFG carefully tracks changes in people's motor behavior in the context of TPEs[62] and responds particularly strongly whenever two people display directly contingent rather than non-contingent movements.[71] Initial data even suggest that the IFG analyzes joint actions in terms of their

overarching goals. Activity increases in the region have been reported, for instance, whenever joint actions serve the completion of an explicit instrumental goal (e.g., carrying a heavy box together) rather than mere affiliation (e.g., hugging each other).[6]

In contrast to the IFG, the role of the IPL during TPE processing is less well understood. Though some studies have reported increased IPL activity in response to TPEs of negative compared to neutral[67] or positive valence[79], others have found no effect of valence[70] or reported the opposite result[61] (i.e., enhanced IPL activity towards positive relative to negative interactions). Interestingly, these conflicting findings may be related to another intriguing observation. At least two studies suggest that IPL activity systematically increases whenever two people enter each other's personal space.[59,82] In other words, the IPL may analyze whether two people get physically so close that they could directly touch (i.e., caress or hit) each other. Given that prior work on valenced TPEs has rarely accounted for potential variations in interpersonal distance across experimental conditions, future investigations are needed to disentangle the relative contributions of variations in valence and interpersonal space on IPL responses.

Future work should also examine the dominant view that the AON enables action understanding through simulation.[54] According to this theory, observers make sense of others' actions by mapping these actions onto their own motor system. This simulation of other people's actions is often declared a hallmark of human social cognition.[109] Yet when observing scenarios that involve two (or more) individuals it is uncertain whose actions (if any) observers would map onto their own motor systems. Initial work on the topic suggests that simulation does not disappear in the face of TPEs (as could be intuitively assumed). Instead, motor evoked potentials (MEPs) indicative of motor stimulation have been found to increase whenever people observe joint rather than individual actions in others[11,12] (e.g., a person throwing a ball to a partner vs. throwing a ball against a wall). However, the work tracking MEPs in response to TPEs has generally portrayed asymmetric encounters between two people in which one (active) individual acts upon another (passive) individual. Given that observers may intuitively take the perspective of the active agent when witnessing such types of TPEs, further research is needed to clarify motor simulation(s) in response to encounters that involve two active partners (e.g., two people greeting each other, carrying a box etc.[110]).

In summary, though the role of the AON during TPE processing requires further investigation, initial findings indicate that the network is highly responsive towards

coordinated actions between multiple individuals, in particular if these actions unfold between people who are in close physical proximity to each other. What seems unclear, at this point, is whether and how TPE processing in the pSTS and the IFG differ from each other. Though these two regions are traditionally discussed in the context of different networks of the social brain, initial findings suggest they may show rather similar response patterns in the context of TPEs (see also Figure 3). Thus, future work is needed to describe commonalities as well as differences between the two in order to enhance our understanding of their unique functional contributions during TPE processing.
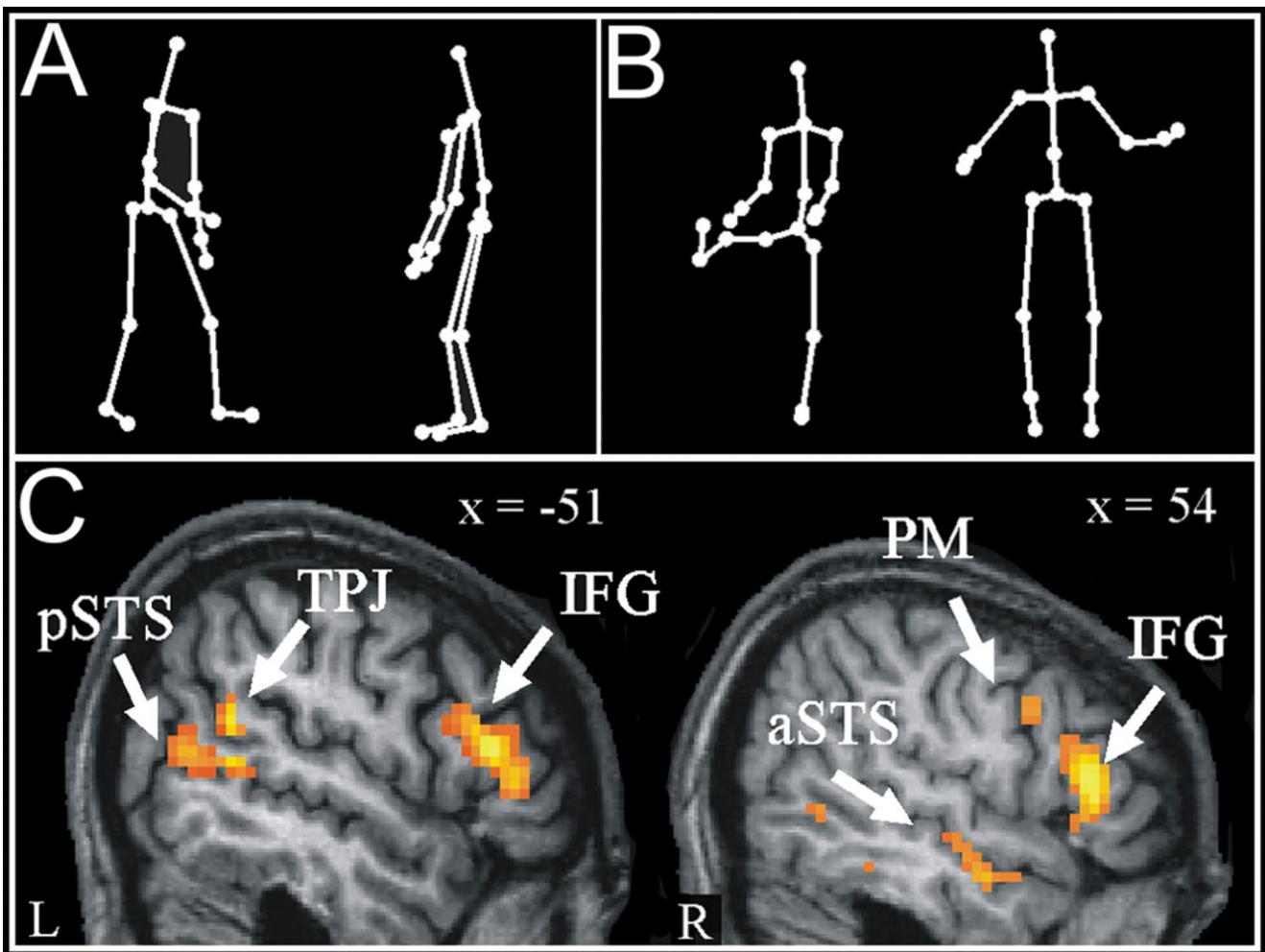

Figure 3. Point-light displays showing person dyads that engage in joint actions (Panel A) versus independent actions (B). To help to distinguish the human form in this illustration, the dots have been linked by full lines. Panel C shows significant activations in response to watching dyads that engage in joint rather than independent actions in the bilateral inferior frontal gyrus (IFG), the right premotor cortex (PM), the right anterior superior temporal sulcus (aSTS), the left posterior superior temporal sulcus (pSTS) and the left temporoparietal junction (TPJ). (Source: ref 58. Image distributed under the Creative Commons Attribution License).

TPE processing in the Mentalizing Network (MTN)

When trying to understand the actions of others, humans frequently rely on attributing invisible mental states (such as desires, motives, intentions, or beliefs) to them. These attributions are widely referred to as 'mentalizing' and hundreds of neuroimaging studies have explored their underlying neural substrates[56]. Collectively, this work supports the notion of a core network for mentalizing that is activated across a wide range of stimuli and tasks.[111] Its key nodes (see Figure 2C) include the ventral and dorsal medial prefrontal cortex (VMPFC and dMPFC), the temporoparietal junction (TPJ), the precuneus (PrC), and the anterior temporal lobe (aTL). For some researchers, also the amygdala – a brain region that is structurally and functionally connected with many nodes of the MTN – forms part of the network (but see ref 112 for counter-evidence).

Importantly, in the context of TPE processing, the MTN (including the amygdala) is certainly the network that has attracted most scientific attention. Numerous studies indicate enhanced activity across all nodes of the MTN towards TPEs compared to non-social controls[74-76,95] as well as compared to portrayals of isolated individuals.[59,64,67,77,78,95] Based on these findings, there remains little doubt that the MTN is particularly tuned towards analyzing the mental states that underlie encounters between multiple individuals. There is further evidence that activity in the network is strongly modulated by variations in TPEs that concern people's actions and/or social relationships. The PrC, in particular, but also the DMPFC and the aTL, have repeatedly been found to distinguish between person dyads that engage in joint rather than in independent actions.[57,58,66,69,78] Further findings that emerge across different studies are less common. For example, comparing TPEs of positive and negative valence has mainly revealed contradictory results. Even though at least three studies have found enhanced DMPFC activity in response to negative compared to neutral[67] or positive[61,70] encounters, others have reported null-findings[79] or the opposite[113] result. Very recent work on the detection of intentional harm in human encounters indicates, however, that simply distinguishing between positive and negative encounters may not suffice to decipher meaningful neural responses. Specifically, converging evidence from fMRI studies[63,80] and intracranial recordings[114] suggests that the amygdala rapidly responds to the occurrence of intentional harm in interpersonal encounters. Importantly, this change in activity is found even when different types of harmful interactions get contrasted, suggesting that it does not reflect a simple valence effect, but the detection of harm that was intentionally caused.[114]

The finding reminds us that the MTN is generally believed to analyze other people's actions in terms of their intentions, desires, and beliefs. Yet very few studies have explicitly manipulated these mental states in the context of TPEs. Furthermore, of those that tried several may be affected by confounds as their experimental conditions have rarely been matched in terms of lower-level perceptual attributes (e.g., the frequency of direct touch between individuals) and/or action properties (e.g., behavior valence). Nevertheless, we want to briefly mention some of the work in order to highlight pivotal dimensions along which TPEs can differ once internal mental states are taken into consideration. For instance, early work on human encounters revealed that interactions between people who differ in terms of their authority (e.g., between a boss and an employee) compared to encounters in which such a difference is absent (e.g., between lovers or siblings) elicited enhanced activity in the aTL.[64] Further work indicates that encounters that primarily serve people's affiliative needs, rather than a well-defined instrumental goal, elicit enhanced activity in the VMPFC and DMPFC.[6] Additionally, there is initial evidence that witnessing changes in a person's body posture that are accompanied by mental changes (e.g., a dad looking up from his newspaper to learn that his son got a bad grade) enhances activity in the aTL, DMPFC and bilateral TPJ.[62]

The latter result deserves particular attention. It has previously been proposed that only the right TPJ is selectively recruited for the attribution of mental states.[115] In context of TPEs, however, TPJ recruitment is often observed bilaterally and/or specifically in the left hemisphere.[57,58,62,76,95] In consequence, it has been proposed that the left TPJ is particularly involved in understanding multiple intentions simultaneously.[76] In support of this idea, a recent study[81] which compared neural responses towards human-human interactions (HHI) with interactions that involved only a single intentional agent (i.e., human-robot interactions, HRI) revealed significantly enhanced activity towards HHI in exactly one region of the MTN – the left TPJ (see Figure 4). This finding strongly confirms the region's pivotal role in analyzing interacting minds. Intriguingly, the reverse contrast revealed enhanced activity in the VMPFC and the precuneus, illustrating that various nodes of the MTN can respond very differently to the same set of TPEs.

Taken together, as is the case for the PPN and AON, the role of the MTN during TPE processing requires further scientific scrutiny. Initial findings indicate, however, that several nodes of the MTN are particularly responsive towards the mental states of co-occurring individuals. To examine this claim more systematically, future work should manipulate

observers' impressions about the mental states of interaction partners more directly in order to understand their neural consequences.[62,69]
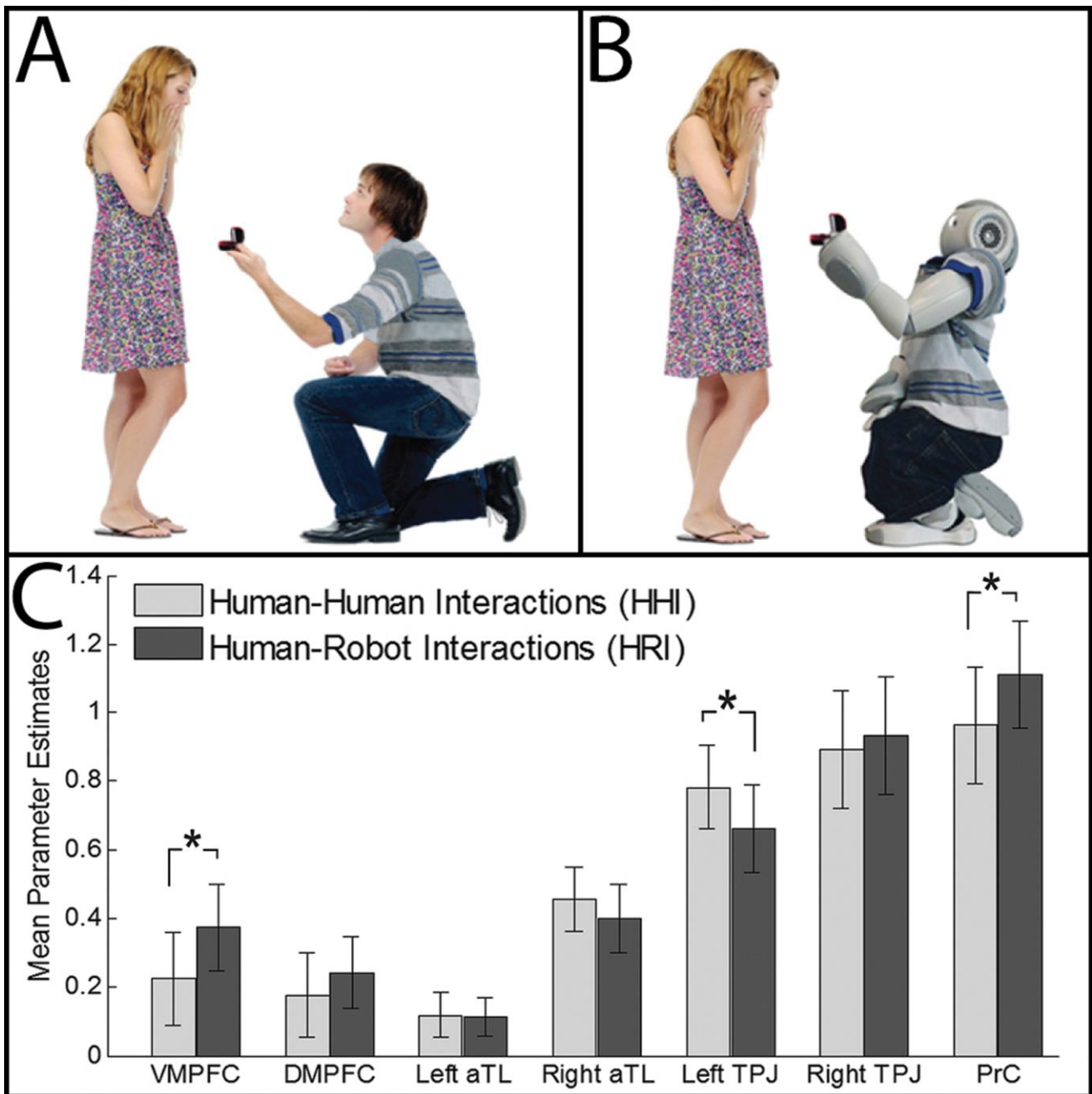

Figure 4. Static color images showing human-human interactions (Panel A) and comparable human-robot interactions (Panel B). Panel C shows brain activity throughout the mentalizing network in response to these two different types of interactions. The network, as localized in this study, includes in the VMPFC (ventromedial prefrontal cortex), DMPFC (dorsomedial prefrontal cortex), aTL (anterior temporal lobe), TPJ (temporoparietal junction), and PrC (precuneus). Intriguingly, only the left TPJ shows a selective activity increase for human-human interactions (Source: ref 81. Copyright 2015 Elsevier, reprinted with permission).

Findings in Clinical Populations

Accumulating evidence indicates that people with disorders that affect social cognition, such as autism[116] and amyotrophic lateral sclerosis[117] are less accurate at forming encounter-based impressions than typical participants. These findings suggest that differences in people's encounter-based impressions may capture socio-cognitive deficits that accompany various psychological and neurological disorders. Such findings are noteworthy, considering there has been little progress to capture these deficits with traditional person perception tasks. Despite twenty years of research, for instance, it remains uncertain which aspects of face perception (if any) are altered in autism[118], including whether the recognition of basic facial expressions of emotions is disturbed.[119]. As such, probing typical and atypical social-cognitive functioning with TPEs promises to provide a particularly fertile avenue for future research.[120]

Indeed, behavioral paradigms which asked participants to view and assess both static and dynamic encounters between others have already been used to better define the nature of socio-cognitive differences in neurodevelopmental disorders such as autism spectrum disorder[116,121,122], Fragile X syndrome[123], Williams Syndrome[124,125], and schizophrenia.[126] Such paradigms have also proven their utility in helping to demonstrate specific social perception and cognition deficits in a few neurodegenerative disorders, including Amyotrophic Lateral Sclerosis[117] as well as Alzheimer's Disease and Frontotemporal Dementia.[127] In combination, these studies suggest that our understanding of a wide range of clinical disorders may ultimately improve by being able to test disorder-specific variations in people-watching.

Despite this potential clinical benefit, few neuroimaging studies have studied TPE processing in clinical populations. Those that have, however, have generally focused on differences between the clinical group and a control group in the MTN. Changes in MTN activity during TPE exposure, for instance, have been demonstrated in schizophrenia[57,78], forensic psychopathy[113], and post-traumatic stress disorder.[70] Most recently, however, it has been shown that idiosyncratic patterns of brain activity during TPE processing throughout the brain can signal substantial deficits in social-cognitive functioning. Specifically, participants with autism spectrum disorder as well as age- and IQ-matched controls were asked to freely watch a TV episode that depicted "constant social interactions that often required perceiving and interpreting subtle, rapidly changing, nuances of facial expressions, body language, and dialogue" (p. 5839) while undergoing fMRI.[128] This task revealed less consistent correlations

in neural activity patterns during TV consumption across individuals of the autism group, an effect that depended mainly on the activity seen in five individuals in this group. Additional analyses demonstrated that these five individuals also struggled more so than the remaining participants with understanding the interpersonal motives and intentions of the depicted TV characters.

To the best of our knowledge, the above finding is the first to directly link neural abnormalities during TPE processing to specific socio-cognitive deficits. Based on this seminal result, it seems rather uncontroversial to assume that gaining a better understanding of the brain regions involved in perceiving and understanding TPEs will aid researchers in their attempt to understand common socio-cognitive deficits that characterize numerous clinical disorders. At the same time, pinpointing differences in brain activity and/or structure that accompany deficits (or strengths!) in social perception and cognition promise to also directly inform our learning about the typical structure and function of the social brain.

Concluding remarks and future directions

To conclude, people-watching entails the continuous analysis of multiple human faces, bodies, and movements in order to understand the course and purpose of social interactions and relations. By describing what is currently known about the neural effects of people-watching in three central networks of the social brain, we aimed to emphasize that making sense of multiple individuals frequently differs and goes beyond forming impressions about isolated individuals. While studying lots of different portrayals of human encounters in the past has elucidated the wide range of visual signals observers can use to form impressions about others, we advocate for a more systematic (i.e., theory-driven and data-driven) approach in future research.

Moving forward, we believe, the field must aspire to use naturalistic portrayals of human encounters more frequently in order to understand the relevance and interplay of various visual markers that determine the type of impressions observers form when witnessing others[95]. At the same time, however, the field should also carefully manipulate these markers in order to learn how exactly they affect observers' psychological and neural responses towards TPEs[66] and control the tasks that people perform during TPE exposure[65] in order to delineate which patterns of neural activity generalize across, or depend on, certain kinds of perceptual, action, or social judgments.

We are further convinced that future research could profit strongly from the use of well-established localizer tasks to better understand how various types of TPEs recruit previously defined brain networks (such as the PPN, AON, and MTN). By localizing brain regions of prior theoretical relevance, neuroscientists could not only determine to which degree the observation and evaluation of TPEs relies on well-known neural structures, but also whether their processing engages additional neutral substrates.[81] For instance, numerous papers as discussed in this review seem to have come across unexpected, yet prevalent activation of the middle temporal gyrus during TPE exposure.[10,58,67,68,78,79] This finding acts as an important reminder that observing and understanding human encounters may recruit neural substrates which the current review has failed to address.

In acknowledgement of this concern, the development of new localizer tasks which capture the most common brain activity during TPE processing in a standardized manner may also be helpful. Bespoke localizer tasks can certainly facilitate the comparison of neuroscientific findings across individuals and studies[129], yet their purpose must be clearly defined. For those interested in learning more about the perceptual encoding of TPEs, for instance, a simple n-back detection task resembling a standard face-localizer[130], but presenting images of isolated person targets and images of (interacting and non-interacting) person dyads and their scrambled controls, may be a good starting point. By contrast, for those interested in examining encounter-based inferences (e.g., social impressions) a categorization task that prompts participants to explicitly classify the same set of person dyads according to their social qualities (e.g., low vs. high dyadic intimacy) as well as action properties (e.g., joint vs. independent actions) may be more fitting.[83] As such, the development of useful TPE localizers poses an interesting challenge for the field.

Equally relevant is the challenge of applying ongoing methodological advances in the neurosciences to the study of encounter-based impressions.[131] In terms of fMRI-based research, for example, it seems crucial to expand the repertoire of analyses beyond standard approaches (such as univariate whole-brain contrasts). Additional methods, such as effective connectivity analysis[120] or multivariate pattern analysis[92] are likely to reveal not only which brain regions are co-active during a given task, but also which of these regions form functional networks by directly aligning their activity and/or representations over time. Furthermore, inter-subject correlation methods promise to be particularly important in discovering developmental or clinical changes in encounter-based impressions. By being able to assess how idiosyncratic an individual's brain response during TPE exposure actually

is[128], such methods promise to lend themselves well to answering pressing questions like "how adult-like is a person's brain response at a specific age", "how typical is a person's brain response compared to a relevant control group" and "how predictive of specific social skills is a person's brain response".

Finally, it is important to keep in mind that much prior work on TPEs, as well as this review, has focused almost exclusively on the processing of human encounters between two individuals. While we believe that understanding the psychological and neural processes involved in the processing of person dyads is of particular importance in the context of human social cognition[132], forming impressions of even larger human gatherings also deserves further investigation.[133-135] By establishing whether systematic differences in brain activity exist based on observing social units of different sizes, important boundary conditions for the different neural networks (or their interplay) may be discovered.

In conclusion, this review suggests that by learning more about how the human brain transforms mere perceptual signals of multiple individuals into far-reaching impressions about them, we may not only develop a neuroscientific framework of people-watching that helps to advance psychological insights on the topic, but also improve our ability to assess and predict social-cognitive deficits in numerous psychological and neurological disorders. It is this promise, after all, that turns a seemingly mundane everyday activity like people-watching into a fascinating topic of multidisciplinary inquiry.

REFERENCES

1. Maclean, E.L., L.J. Matthews, B.A. Hare, C.L. Nunn, R.C. Anderson et al. 2012. How does cognition evolve? Phylogenetic comparative psychology. *Anim. Cogn.* **15:** 223-238.
2. Downing, P.E., & M.V. Peelen. 2016. Body selectivity in occipitotemporal cortex: Causal evidence. *Neuropsychologia*, **83:** 138-148.
3. Yovel, G. 2016. Neural and cognitive face-selective markers: An integrative review. *Neuropsychologia*, **83:** 5-13.
4. Quadflieg, S., & I. Penton-Voak. in press. The emerging science of people-watching: Forming impressions from third-party encounters. *Curr. Dir. in Psychol. Sci.*
5. Burgoon, J.K., D.B. Buller, J.L. Hale, & M.A. de Turck. 1984. Relational messages associated with nonverbal behaviors. *Hum. Commun. Res.* **10:** 351-378.
6. Canessa, N., F. Alemanno, F. Riva, A. Zani, A.M. Proverbio, N. Mannara, D. Perani, & S.F. Cappa. 2012. The neural bases of social intention understanding: The role of interaction goals. *PLoS One* **7:** e42347.
7. Proverbio, A. M., F. Riva, L. Paganelli, S.F. Cappa, N. Canessa, D. Perani, & A. Zani. 2011. Neural coding of cooperative and affective human interactions: 150 ms to code the action's purpose. *PLoS One* **6:** e22026.
8. Yoder, K.J., & J. Decety. 2014. Spatiotemporal neural dynamics of moral judgment: A high-density ERP study. *Neuropsychologia* **60:** 39-45.
9. Petrini, K., L. Piwek, F. Crabbe, F.E. Pollick, & S. Garrod. 2014. Look at those two!: The precuneus' role in unattended third-person perspective of social interactions. *Hum. Brain Mapp.* **35:** 5190-5203.
10. Quadflieg, S., F. Gentile, & B. Rossion. 2015. The neural basis of perceiving person interactions. *Cortex* **70:** 5-20.
11. Aihara, T., S. Yamamoto, H. Mori, H., K. Kushiro, & S. Uehara. 2015. Observation of interactive behaviour increases corticospinal excitability in humans: A transcranial magnetic stimulation study. *Brain Cogn.* **100:** 1-6.
12. Bucchioni, G., A. Cavallo, D. Ippolito, G. Marton, & U. Castiello. 2013. Corticospinal excitability during the observation of social behavior. *Brain Cogn* **81:** 176-182.
13. Chakroff, A., J. Dungan, J. Koster-Hale, A. Brown, R. Saxe, & L. Young. 2016. When minds matter for moral judgment: intent information is neurally encoded for harmful but not impure acts. *Soc. Cogn. Affect. Neurosci.* **11:** 476-484.
14. Rapp, A.M., K. Langohr, D.E. Mutschler, S. Klingberg, B. Wild, & M. Erb. 2013. Isn't it ironic? Neural correlates of irony comprehension in schizophrenia. *PLoS One* **8:** e74224.
15. Castelli, F., F.G. Happé, U. Frith, & C.D. Frith. 2000. Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage* **12:** 314-325.
16. Gao, T., B.J. Scholl, & G. McCarthy. 2012. Dissociating the detection of intentionality from animacy in the right posterior superior temporal sulcus. *J. Neurosci.* **32:** 14276-14280.
17. Jack, A., & K.A. Pelphrey. 2015. Neural correlates of animacy attribution include neocerebellum in healthy adults *Cereb. Cortex* **25:** 4240-4247.
18. Heider, F., & M. Simmel. 1944. An experimental study of apparent behavior. *Am. J. Psychiatry* **57:** 243-259.
19. Macrae, C.N., & S. Quadflieg. 2010. Perceiving people. In S. Fiske, D.T. Gilbert, & G. Lindzey. (Eds.) *The Handbook of Social Psychology* (5th ed., pp. 428-463). New York: McGraw-Hill.

20. Shiffrar, M. 2011. People watching: visual, motor, and social processes in the perception of human movement. *Wiley Interdiscip. Rev. Cogn. Sci* **2:** 68-78.
21. Duncan, B.L. 1976. Differential social perception and attribution of intergroup violence: Testing the lower limits of stereotyping Blacks. *J. Pers. Soc. Psychol.* **34:** 590-598.
22. Trout, D.L., & H.M. Rosenfeld 1980. The effect of postural lean and body congruent on the judgment of psychotherapeutic rapport. *J Nonverbal Behav.* **4:** 176-190.
23. James, J. 1953. The distribution of free-forming small group size. *Am. Sociol. Rev.* **18:** 569‑570.
24. Sorokin, P.A., & C.Q. Berger. 1939. *Time-budgets of human behavior.* Cambridge, MA: Harvard University Press.
25. Wheeler, L., & J. Nezlek. 1977. Sex differences in social participation. *J Pers. Soc. Psychol.* **35:** 742-754.
26. Knowles, E.S. 1973. Boundaries around group interaction: The effect of group size and member status on boundary permeability. *J. Pers. Soc. Psychol.* **26:** 327-331.
27. Knowles, E.S. 2015. Spatial behavior of individuals and groups. In P.B. Paulus (ed.) *Psychology of group influence* (2nd edition, pp. 53-86). New York: Psychology Press.
28. Augusti, E.-M., A. Melinder, & G. Gredebäck. 2010. Look who's talking: pre-verbal infants' perception of face-to-face and back-to-back social interactions. *Front. Psychol.* **1:** 161.
29. Neri, P., J.Y. Luu, & D.M. Levi. 2006. Meaningful interactions can enhance visual discrimination of human agents. *Nat. Neurosci.* **9:** 1186-1192.
30. Rhodes, M., & L. Chalik. 2013. Social categories as markers of intrinsic interpersonal obligations. *Psychol. Sci.* **24:**999-1006.
31. Hinde, R.A. 1976. Interactions, relationships, and social structure. *Man* **11:** 1-17.
32. Bernieri, F.J., J.S. Gillis, J.M. Davis, & J.E. Grahe. 1996. Dyad rapport and the accuracy of its judgment across situations: A Lens Model analysis. *J. Pers. Soc. Psychol.* **71:** 110-129.
33. Pryor, J.B., G.D. Reeder, & A.E. Monroe. 2012. The infection of bad company: stigma by association. *J. Pers. Soc. Psychol.* **102:** 224-241.
34. Manera, V., M. Del Giudice, B.G. Bara, K. Verfailie, & C. Becchio. 2011. The second-agent effect: communicative gestures increase the likelihood of perceiving a second agent. *PLoS One* **6:** e22650.
35. Fawcett, C., & G. Gredebäck. 2013. Infants use social context to bind actions into a collaborative sequence. *Dev. Sci.* **16:** 841-849.
36. Fiedler, K., & W. Schenk. 2001. Spontaneous inferences from pictorially presented behaviors. *Pers. Soc. Psychol. Bull.* **27:** 1533-1546.
37. Floyd, K., & L.A. Erbert. 2003. Relational message interpretations of nonverbal matching behavior: An application of the social meaning model. *J. Soc. Psychol.* **143:** 581-597.
38. Grahe, J.E., & F.J. Bernieri. 1999. The importance of nonverbal cues in judging rapport. *J. Nonverbal Behav.* **23:** 253-269.
39. Costanzo, M., & D. Archer. 1989. Interpreting the expressive behavior of others: the interpersonal perception task. *J. Nonverbal Behav.* **13:** 225-244.
40. Schmid Mast, M., & J.A. Hall. 2004. Who is the boss and who is not? Accuracy of judging status. *J. Nonverbal Behav.* **28:** 145-165.
41. Bryant, G.A., D.M.T. Fessler, R., Fusaroli, E. Clint, L. Aarøe, C.L. Apicella et al. 2016. Detecting colaughter across 24 societies. *Proc. Natl. Acad. Sci. USA* ***113:*** 4682-4687.
42. Seyfarth, R.M. & D.L. Cheney. 2013. Affiliation, empathy, and the origins of theory of mind. *Proc. Natl. Acad. Sci. USA* **110:** 10349–10356.

43. Brent, L.J.N., S.W.C. Chang, J.-F. Gariépy, & M.L. Platt. 2014. The neuroethology of friendship. *Ann. N. Y. Acad. Sci.* **1316:** 1-17.
44. Milinski, M. 2016. Reputation, a universal currency for human social interactions. *Philos. Trans. R. Soc. B* **371:** 20150100.
45. Fawcett, C., & U. Liszkowski. 2012. Observation and initiation of join action in infants. *Child Dev.* **83:** 434-441.
46. Matheson, H., Moore, C., & N. Akhtar. 2013. The development of social learning in interactive and observational contexts. *J. Exp. Child Psychol.* **114:** 161-172.
47. Mazziotta, A., A., Mummenday, & S.C. Wright. 2011. Vicarious intergroup contact effects: Applying social-cognitive theory to intergroup contact research. *Group Process. Intergr. Relat.* **14:** 255-274.
48. Shimpi, P.M., N. Akhtar, & C. Moore. 2013. Toddlers' imitative learning in interactive and observational contexts: The role of age and familiarity of the model. *J. Exp. Child Psychol.* **116:** 309-323.
49. Willingham, D.T., & E.W. Dunn. 2003. What neuroimaging and brain localization can do, cannot do, and should not do for social psychology. *J. Pers. Soc. Psychol.* **85:** 662-671.
50. Yang, D.Y.-J., G. Rosenblau, C. Keifer, & K.A. Pelphrey. 2015. An integrative neural model of social perception, action observation, and theory of mind. *Neurosci. Biobehav. Rev.* **51:** 263-275.
51. Downing, P.E., & M.V. Peelen. 2016. Body selectivity in occipitotemporal cortex: Causal evidence. *Neuropsychologia* **83:** 138-148.
52. Yovel, G. 2016. Neural and cognitive face-selective markers: An integrative review. *Neuropsychologia* **83:** 5-13.
53. Gardner, T., N. Goulden, & E.S. Cross. 2015. Dynamic modulation of the action observation network by movement familiarity. *J. Neurosci.* **35:** 1561-1572.
54. Rizzolatti, G., L. Fogassi, & V. Gallese. 2001. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci.* **2:** 661-670.
55. Frith, U. & C.D. Frith. 2003. Development and neurophysiology of mentalizing. *Philos. Trans. R. Soc. B* **358:** 459-473.
56. Schurz, M., J. Radua, M. Aichhorn, F. Richlan, & J. Perner. 2014. Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neurosci. Biobehav. Rev.* **42:** 9-34.
57. Bakasch, B., B. Straube, M. Pyka, F. Klöhn-Saghatolislam, M.J. Müller, T.T.J. Kircher, et al. 2013. Hyperintentionality during automatic perception of naturalistic cooperative behavior in patients with schizophrenia. *Soc. Neurosci.* **8:** 489-504.
58. Centelles, L., C. Assaiante, B. Nazarian, J.-L. Anton, & C. Schmitz. 2011. Recruitment of both the mirror and mentalizing networks when observing social interactions depicted by point-lights: A neuroimaging study. *PloS One* **6:** e15749.
59. Dolcos, S., K. Sung, J.J. Argo, S. Flor-Henry, & F. Dolcos. 2012. The power of a handshake: Neural correlates of evaluative judgments in observed social interactions. *J. Cogn. Neurosci.* **24:** 2292-2305.
60. Kujala, M.V., S. Carlson, & R. Hari. 2012. Engagement of amygdala in third-person view of face-to-face interaction. *Hum. Brain Mapp.* **33:** 1753-1762.
61. Deuse, L., L.M. Rademacher, L. Winkler, R.T. Schultz, G. Gründer, & S.E. Lammertz. 2016. Neural correlates of naturalistic social cognition: brain-behavior relationships in healthy adults. *Soc. Cogn. Affect. Neurosci.* **11:** 1741-1751.
62. Hooker, C.I., S.C. Verosky, L.T. Germine, R.T. Knight, & M. D'Esposito. 2010. Neural activity during social signal perception correlates with self-reported empathy. *Brain Res.* **1308:** 100-113.

63. Molenberghs, P., Gapp, J., Wang, B., Louis, W. R., & Decety, J. (2016). Increased moral sensitivity for outgroup perpetrators harming ingroup members. *Cereb. Cortex* **26:** 225-233.

64. Iacobini, M., M.D. Lieberman, B.J. Knowlton, I. Molnar-Szakacs, M. Moritz, C.J. Throop, & A.P. Fiske. 2004. Watching social interactions produces dorsomedial prefrontal and medial parietal BOLD fMRI signal increases compared to a resting baseline. *Neuroimage* **21:** 1167-1173.

65. Mason, M., J.C. Magee, & S.T. Fiske. 2014. Neural substrates of social status inference: Roles of medial prefrontal cortex and superior temporal cortex. *J. Cogn. Neurosci.* **26:** 1131-1140.

66. Pierno, A.C., C. Becchio, L. Turella, F. Tubaldi, & U. Castiello. 2008. Observing social interactions: The effect of gaze. *Soc. Neurosci.* **3:** 51-59.

67. Krämer, U.M., B. Mohammadi, N. Doñamayor, A. Samii, & T.F. Münte. 2010. Emotional and cognitive aspects of empathy and their relation to social cognition – an fMRI-study. *Brain Res.* **1311:** 110-120.

68. Wang, T., L. Mo, C. Mo, L.H. Tan, J.S. Cant, L. Zhong, & G. Cupchick. 2015. Is moral beauty different from facial beauty? Evidence from an fMRI study. *Soc. Cogn. Affect. Neurosci.* **10:** 814-823.

69. Eskenazi, T., S.-A. Rueschemeyer, F.P. de Lange, G. Knoblich, & N. Sebanz. 2015. Neural correlates of observing joint actions with shared intentions. *Cortex* **70:** 90-100.

70. Moser, D.A., T. Aue, F. Suardi, H. Kutlikova, M.I. Cordero, A.S. Rossignol, et al. 2015. Violence-related PTSD and neural activation when seeing emotionally charged male-female interactions. *Soc.I Cogn. Affect. Neurosci.* **10:** 645-653.

71. Georgescu, A.L., B. Kuzmanovic, N.S. Santos, R. Tepest, G. Bente, M. Tittgemeyer, & K. Vogeley. 2014. Perceiving nonverbal behaviour: Neural correlates of processing movement fluency and contingency in dyadic interactions. *Hum. Brain Mapp. 35:* 1362-1378.

72. Ulloa, J.L., A. Puce, L. Hugueville, & N. George. 2014. Sustained neural activity to gaze and emotion perception in dynamic social scenes. *Soc. Cogn. Affect. Neurosci.* **9:** 350-357.

73. Shibata, H., T. Inui, & K. Ogawa. 2013. Role of the dorsolateral prefrontal cortex in recognizing hand actions performed in social contexts: a functional MRI study. *Neuroreport* **24:** 803-807.

74. Chen, C., J. Decety, P.C. Huang, C.-Y. Chen, & Y. Cheng. 2016. Testosterone administration in females modulates moral judgment and patterns of brain activation and functional connectivity. *Hum. Brain Mapp.* **37:** 3417-3430.

75. Vrtička, P., D. Sander, & P. Vuilleumier. 2011. Effects of emotion regulation strategy on brain responses to valence and social content of visual scenes. *Neuropsychologia* **49:** 1067-1082.

76. Ciaramidaro, A., M. Adenzato, I. Enrici, S. Erk, L. Pia, B.G. Bara, & H. Walter. 2007. The intentional network: How the brain reads varieties of intentions. *Neuropsychologia* **45:** 3105-3113.

77. Walter, H., M. Adenzato, A. Ciaramidaro, I. Enrici, L. Pia, & B.G. Bara. (2004). Understanding intentions in social interaction: The role of the anterior paracingulate cortex. *J. Cogn. Neurosci.* **16:** 1854-1863.

78. Leube, D., B. Straube, A. Green, I. Blümel, S. Prinz, P. Schlotterbeck, & T. Kircher. 2012. A possible brain network for [the] representation of cooperative behavior and its implications for the psychopathology of schizophrenia. *Neuropsychobiology* **66:** 24-32.

79. Yoder, K.J., & J. Decety. 2014. The good, the bad, and the just: Justice sensitivity predicts neural response during moral evaluation of actions performed by others. *J. Neurosci.* **34:** 4161-4166.

80. Sinke, C.B.A., B. Sorger, R. Goebel, & B. de Gelder. 2010. Tease or threat? Judging social interactions from bodily expressions. *NeuroImage* **49:** 1717-1727.

81. Wang, Y., & S. Quadflieg. 2015. In our own image? Emotional and neural processing differences when observing human-human versus human-robot interactions. *Soc.l Cogn. Affect. Neurosci.* **10:** 1515-1524.

82. Lloyd, D.M., & C.I. Morrison. 2008. 'Eavesdropping' on social interactions biases threat perception in visuospatial pathways. *Neuropsychologia* **46:** 95-101.

83. Spunt, R.P., & R. Adolphs. 2014. Validating the Why/How contrast for functional MRI studies of Theory of Mind. *Neuroimage* **99:** 301-311.

84. Kanwisher, N., J. McDermott, & M.M. Chun. 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* **17:** 4302-4311.

85. Haxby, J.V., E.A. Hoffman, & M.A. Gobbini. 2000. The distributed human neural system for face perception. *Trends Cogn. Sci.* **4:** 223-233.

86. Kanwisher, N. & G. Yovel. 2006. The fusiform face area: a cortical region specialized for the perception of faces. *Philos. Trans. R. Soc. B* **361:** 2109-2128.

87. Downing, P.E., Y. Jiang, M. Shuman, & N. Kanwisher. 2001. A cortical area selective for visual processing of the human body. *Science* **293:** 2470-2473.

88. Grossman, E., M. Donnelly, R. Price, D. Pickens, V. Morgan, J. Neighbor, & R. Blake. 2000. Brain areas involved in perception of biological motion. *J. Cogn. Neurosci.* **12:** 711-720.

89. Haxby, J.V., & M.I. Gobbini. 2011. Distributed neural systems for face perception. In G. Rhodes, A. Calder, M. Johnson, & J.V. Haxby (Eds.) *The Oxford Handbook of face perception* (pp. 93-110). Oxford University Press.

90. Yovel, G., & A.J.O. O'Toole. 2016. Recognizing people in motion. *Trends Cogn. Sci.* **20:** 383-395.

91. Bernstein, M., J. Oron, B. Sadeh, & G. Yovel. 2014. An integrated face-body representation in the fusiform gyrus but not the lateral occipital cortex. *J. Cogn. Neurosci.* **26:** 2470-2478.

92. Brandman, T., & G. Yovel. 2016. Bodies are represented as wholes rather than the sum of parts in the occipital-temporal cortex. *Cereb. Cortex* **26:** 530-543.

93. Wagner, D.D., W.M. Kelley, & T.F. Heatherton. 2011. Individual differences in the spontaneous recruitment of brain regions supporting mental state understanding when viewing natural social scenes. *Cereb. Cortex* **21:** 2788-2796.

94. Deen, B., K. Koldewyn, N. Kanwisher, & R. Saxe. 2015. Functional Organization of Social Perception and Cognition in the Superior Temporal Sulcus. *Cereb. Cortex* **25:** 4596-4609.

95. Wagner, D.D., W.M. Kelley, J.V. Haxby, & T.F. Heatherton. 2016. The dorsal medial prefrontal cortex responds preferentially to social interactions during natural viewing. *J. Neurosci.* **36:** 6917-6925.

96. Pitcher, D., D.D. Dilks, R. Saxe, C. Triantafyllou, & N. Kanwisher. 2011. Differential selectivity for dynamic versus static information in face selective cortical regions. *Neuroimage* **56:** 2356-2363.

97. Downing, P.E., & M.V. Peelen. 2011. The role of occipitotemporal body-selective regions in person perception. *Cogn. Neurosci.* **2:** 186-203.

98. Kilner, J.M. 2011. More than one pathway to action understanding. *Trends Cogn. Sci.* **15:** 352-357.

99. Lingnau, A., & P.E. Downing. 2015. The lateral occipitotemporal cortex in action. *Trends Cogn. Sci.* **19:** 268-277.

100. Baek, A., J. Wagemans, & H.P. Op de Beeck. 2013. The distributed representation of random and meaningful object pairs in human occipitotemporal cortex: The weighted average as a general rule. *Neuroimage* **70:** 37-47.

101. Baldassano, C., D.M. Beck, & L. Fei-Fei. in press. Human-object interactions are more than the sum of their parts. *Cereb Cortex*

102. Buccino, G., F. Binkofski, G.R. Fink, L. Fadiga, L. Fogassi, V. Gallese, et al. 2001. Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur. J. Neurosci.* **13:** 400-404.

103. Chong, T.T.J., M.A. Williams, R. Cunnington, & J.B. Mattingley. 2008. Selective attention modulates inferior frontal gyrus activity during action observation. *Neuroimage* **40:** 298-307.

104. Saygin, A.P., S.M. Wilson, D.J. Hagler, E. Bates, & M.I. Sereno. 2004. Point-light biological motion perception activates human premotor cortex. *J. Neurosci.* **24:** 6181-6188.

105. Thompson, J. & R. Parasuraman. 2012. Attention, biological motion, and action recognition. *Neuroimage* **59:** 4-13.

106. Caspers, S., K. Zilles, A.R. Laird, & S.B. Eickhoff. 2010. ALE meta-analysis of action observation and imitation in the human brain. *Neuroimage* **50:** 1148–1167.

107. Grèzes, J., & J. Decety. 2001. Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. *Hum. Brain Mapp.* **12:** 1-19.

108. Molenberghs, P., R. Cunnington, & J.B. Mattingley. 2012. Brain regions with mirror properties: a meta-analysis of 125 human fMRI studies. *Neurosci. Biobehav. Rev.* **36:** 341-349.

109. Hasson, U. & C.D. Frith. 2016. Mirroring and beyond: coupled dynamics as a generalized framework for modelling social interactions. *Philos. Trans. R. Soc. B* **371:** 20150366.

110. Donne, C.M., P.E. Enticott, N.J. Rinehart, & P.B. Fitzgerald. 2011. A transcranial magnetic stimulation study of corticospinal excitability during the observation of meaningless, goal-directed, and social behavior. *Neurosci. Lett.* **489:** 57-61.

111. Mar, R.A. 2011. The neural bases of social cognition and story comprehension. *Ann. Rev. Psychol* **62:** 103-134.

112. Spunt, R.P., J.T. Elison, N. Dufour, R. Hurlemann, R. Saxe & R. Adolphs. 2015. Amygdala lesions do not compromise the cortical network for false-belief reasoning. *Proc. Natl. Acad. Sci. USA* **112:** 4827-4832.

113. Decety, J., C. Chen, C.L. Harenski, & K.A. Kiehl. 2015. Socioemotional processing of morally-laden behaviour and their consequences on others in forensic psychopaths. *Hum. Brain Mapp.* **36:** 2015-2026.

114. Hesse, E., E. Mikulan, J. Decety, M. Sigmna, M. del Carmen Garcia, S. Walter, et al. 2016. Early detection of intentional harm in the human amygdala. *Brain* **139:** 54-61.

115. Saxe, R., & A. Wexler. 2005. Making sense of another mind: the role of the right temporoparietal junction. *Neuropsychologia* **43:** 1391-1399.

116. Centelles, L., C. Assaiante, K. Etchegoyhen, M. Bouvard, M., & C. Schmitz. 2013. From action to interaction: Exploring the contribution of body motion cues to social understanding in typical development and in autism spectrum disorders. *J. Aut. Dev. Disord.* **43:** 1140-1150.

117. Cavallo, M., M. Adenzato, S.E. MacPherson, G. Karwig, I. Enrici, & S. Abrahams. 2011. Evidence of social understanding impairment in patients with amyotrophic lateral sclerosis. *PLoS One* **6:** e25948.
118. Weigelt, S., K. Koldewyn, & N. Kanwisher. 2012. Face identity recognition in autism spectrum disorders: a review of behavioral studies. *Neurosci. Biobehav. Rev.* **36:** 1060-1084.
119. Uljarevic, M., & A. Hamilton. 2013. Recognition of emotions in autism: a formal meta-analysis. J Autism Dev Disord **43:** 1517-1526.
120. Henry, J.D., W. von Hippel, P. Molenberghs, T. Lee, & P.S. Sachdev. 2016. Clinical assessment of social cognitive function in neurological disorders. *Nat. Rev. Neurolog.* **12:** 28-39.
121. Bauminger, N., C. Shulman, & G. Agam. 2004. The link between perceptions of self and of social relationships in high-functioning children with autism. *J. Dev. Phys. Disabil. 16*: 193-214.
122. Klin, A., W. Jones, R. Schultz, F. Volkmar, & D. Cohen. 2002. Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Arch Gen Psychiatry* **59:** 809-816.
123. Williams, T.A., M.A. Porter, & R. Langdon. 2013. Viewing social scenes: A visual scan-path study comparing Fragile X Syndrome and Williams Syndrome. *J Aut. Dev. Disord.* **43:** 1880-1894.
124. Dodd, H.F., & M.A. Porter 2011. Interpretation of ambiguous social situations: Evidence for a dissociation between social and physical threat in Williams Syndrome*. J Autism Dev Disord* **41:** 266-274.
125. Riby, D.M., & P.J.B. Hancock. 2008. Viewing it differently: Social scene perception in Williams syndrome and Autism. *Neuropsychologia* **46***: 2855-2860.
126. Okruszek, L., M. Haman, K. Kalinowski, M. Talarowska, C. Becchio, & V. Manera. 2015. Impaired recognition of communicative interactions from biological motion in schizophrenia. *PLoS One* **10:** e0116793.
127. Cavallo, M., I. Enrici, & M. Adenzato. 2011. The comprehension of social situations in a small gorup of patients with Frontotemporal Dementia and Alzheimer's disease. *Acta Neuropsychol* **9:** 167-176.
128. Byrge, L., J. Dubois, J.M. Tyszka, R. Adolphs, & D.P. Kennedy. 2015. Idiosyncratic brain activation patterns are associated with poor social comprehension in autism. *J. Neurosci.* **35:** 5837-5850.
129. Nieto-Castañón, A., & E. Fedorenko. 2012. Subject-specific functional localizers increase sensitivity and functional resolution of multi-subject analyses. *Neuroimage* 63: 1646-1669.
130. Rossion, B., B. Hanseeuw, & L. Dricot. 2012. Defining face perception areas in the human brain: A large-scale factorial fMRI face localizer analysis. *Brain Cogn.* 79: 138-157.
131. Adolphs, R., L. Nummenmaa, A. Todorov, & J.V. Haxby. 2016. Data-driven approaches in the investigation of social perception. *Philos. Trans. R. Soc. B* 371: 20150367.
132. Kelley, H.H., J.G. Holmes, N.L. Kerr, H.T. Reis, C.E. Rusbult, & P.A.M. Van Lange. 2003. *An atlas of interpersonal situations*. New York, NY: Cambridge University Press.
133. Elias, E., M. Dyer, & T.D. Sweeny. in press. Ensemble perception of dynamic emotional groups. *Psychol. Sci.*
134. Jenkins, A.C., D. Dodell-Feder, R. Saxe, & J. Knobe. 2014. The neural bases of directed and spontaneous mental state attributions to group agents. *PLoS One* 9: e105341.

135. Yamanashi Leib, A., J. Fischer, Y. Liu, S. Qiu, L. Robertson, & D. Whitney. 2014. Ensemble crowd perception: a viewpoint-invariant mechanism to represent average crowd identity. *J. Vis.* 14: 26(1-13).