Neural Correlates of Empathy in Humans, and the Need for Animal Models

Keysers, C.; Gazzola, V.

DOI
10.1016/B978-0-12-805397-3.00004-8

Publication date
2018

Document Version
Final published version

Published in
Neuronal Correlates of Empathy

CHAPTER

4

Neural Correlates of Empathy in Humans, and the Need for Animal Models

Christian Keysers[*,**], *Valeria Gazzola*[*,**]

***Netherlands Institute for Neuroscience, an Institute of the Royal Netherlands Academy of Art and Sciences (KNAW), Amsterdam, The Netherlands; **University of Amsterdam, Amsterdam, The Netherlands**

EMPATHY AND ITS COMPONENTS: EMOTIONAL CONTAGION, EMPATHY PROPER, AND SYMPATHY

Empathy often refers to a family of related processes through which we intuitively relate to the inner states of others. Although definitions vary, there is some consensus that one should discriminate at least three different phenomena when it comes to how an individual (human or animal) reacts to the emotions of another individual (Keysers, 2011; de Vignemont & Singer, 2006; Wispé, 1986). We will refer throughout this text to the individual experiencing empathy as a "witness" (rather than an observer, to underline the importance of senses other than vision), and to the person that is the object of the empathy as the "object." The first and simplest phenomenon is often referred to as *emotional contagion*, and occurs whenever a witness's emotional states come to resemble those of the object. For this simplest form to occur, it is not necessary for the witness to clearly attribute her triggered internal feelings to the feelings of the object. A prototypical example is when a newborn baby starts crying when witnessing other babies cry. The second phenomenon, *empathy proper*, occurs when the witness experiences emotional states that resemble those of the object and is aware that these states

Neuronal Correlates of Empathy. http://dx.doi.org/10.1016/B978-0-12-805397-3.00004-8

are experienced on behalf of the object. In both emotional contagion and empathy, a witness' triggered emotional states thus resemble those of the object of empathy, such that witnessing a person demonstrating distress triggers distress in the witness. This is where the third phenomenon, *sympathy*, differs: sympathy occurs when the witness no longer directly experiences emotions resembling those of the object, but rather prosocial emotions that are appropriate to help the object. Sympathy for a person in distress is then no longer distress, but a combination of distress and a "warm-hearted feeling" that motivates the witness to console and help the object.

It is generally believed that these three phenomena are causally related, such that overt signs of an object's emotions first trigger that emotion in a witness through emotional contagion. If the meta-cognitive apparatus of the witness enables her/him to do so, this contaged emotion is then attributed to the object of empathy, triggering an additional layer of empathy proper. If the witness is so disposed, this can then lead to the prosocial motivation called sympathy.

This trilayered distinction finds its routes in human psychology, where these phenomena are thought to appear in succession over the course of development, with emotional contagion being present soon after birth, while empathy proper and sympathy appear later, when children develop the ability to represent the self and others separately, and are able to regulate their own emotions well enough to help others (Decety & Jackson, 2004). Directly applying these distinctions in a cross-species approach is challenging: although rodents show signs of altering their own emotions in reaction to those of others (Atsak et al., 2011; Panksepp & Panksepp, 2013) and also demonstrate a motivation to help animals in needs (Ben-Ami Bartal, Decety, & Mason, 2011; Greene, 1969; Rice & Gainer, 1962), it is not obvious how one would establish whether rats are aware that the emotion they experience is that of another rat. Whether they experience emotional contagion, empathy proper, or sympathy thus remains difficult to assess. We therefore recommend using the term empathy and study its neural basis initially in its widest sense to gain traction on how the emotions of an object affect those of the witness, and then later attempt to distinguish emotional contagion, empathy proper, and sympathy.

These three layers of social transmission of affect exist in the context of a number of related phenomena. Some of them relate to how we perceive the actions of others, as well as their somatosensory states. Some are related to more explicit cognitive processes and have been referred to as mentalizing. Before focusing on emotional empathy, we will therefore briefly touch on this context.

RELATED PHENOMENA: MENTALIZING, SOMATOSENSORY, AND MOTOR EMPATHY

Mentalizing refers to people's ability to consciously think about the beliefs, desires, and intensions of others. Mentalizing is sometimes also referred to as "cognitive empathy." If we see a friend in great distress, and start to mentally deliberate that this is probably because he did not know that an important grant deadline had been moved, we engage in mentalizing. If the same friend simply makes us feel agitated without us engaging in the attribution of specific beliefs, we experience emotional contagion or empathy (depending on how much we attribute our distress to him). Clearly, these phenomena can—and often will—complement each other. However, because of space constrains, we will not be able to touch explicitly on the relation between the neural correlates of mentalizing and other forms of empathy, but a discussion of this relation can be found in Keysers and Gazzola (2007).

Somatosensory and motor empathy, on the other hand, refers to the fact that while witnessing other people's movements, we can feel ourselves into their movements as if we were moving in similar ways, a phenomenon sometimes also called kinesthetic empathy, particularly in the scientific study of the esthetics of dance (Jola, Ehrenberg, & Reynolds, 2012). Because we cannot telepathically feel the emotions of others, perceiving them must be based on the object's behavior: her bodily movements, facial expressions, vocalizations, the activity of glands (tears, sweat, pheromones) and, in humans, linguistic behavior—or the conspicuous lack of any of those. As such, mechanisms that allow us to process the actions of others are highly relevant to processing their emotions, which is why we will start by reviewing what we know about the neural correlates of kinesthetic empathy before addressing the neural correlates of emotional contagion in humans.

Our understanding of how the brain processes the actions of others has been strongly influenced by the discovery of mirror neurons in the macaque monkey (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996). These motor neurons, first described in the ventral premotor cortex of the macaque monkey, respond both when a monkey performs a goal-directed action, for instance breaking the shell of a peanut, and when witnessing a similar action performed by others, be it by only seeing someone performing a silent version of this action or by only hearing the cracking of the shell or by hearing and seeing the action (Keysers et al., 2003; Kohler et al., 2002). Because electrostimulating the premotor cortex of the monkey triggers the performance of similar actions (Graziano, 2016), this suggest that witnessing the actions of others triggers a motor representation fit

for performing a similar action in the witness. About 10% of the motor neurons in the ventral premotor cortex of the macaque have this mirror property, while the remaining 90% do not respond to the sight or sound of actions. Because the firing of premotor neurons is associated not with simple movements of muscles, but with complex, goal-directed actions such as grasping or tearing, and often do so independently of whether the action is performed with the right or the left hand, or even with the mouth (Gallese et al., 1996), this suggests that what is encoded in these neurons is not simple movements, such as flexing the biceps, but a higher-level action policy of what the action is to achieve (e.g., cracking a peanut's shell). This goal-directed coding becomes particularly apparent in mirror neurons that respond to the sound of actions: if a neuron responds when the monkey himself cracks a peanut, as well as when listening to the sound of a cracking peanut shell, what is common is the goal (cracking the shell), not the means—as the body movements achieving the goal are not apparent from the sound (Keysers, 2011).

A decade later, similar neurons were reported in the rostral inferior parietal cortex (Fogassi et al., 2005) of the monkey, a region that has direct connections to the ventral premotor cortex, which is also involved in the motor control of goal-directed actions.

In humans, rigorous functional magnetic resonance imaging studies have mapped all brain regions that have voxels that are activated both while executing and witnessing the actions of others. For instance we had participants view a number of complex goal-directed hand actions and execute such actions while their brain activity was measured using fMRI (Gazzola & Keysers, 2009). Voxels were then identified as being shared between action observation and execution, and hence to be potentially involved in kinesthetic empathy, on an individual participant level, without the spatial smoothing that is normally used in fMRI, revealing a network of brain regions active both during action observation and execution. This includes regions traditionally associated with motor control, including the posterior parietal cortex, the dorsal and ventral premotor cortex, the supplementary motor cortex, and the cerebellum. That the ventral premotor and posterior parietal cortex had this property is in line with the monkey studies showing mirror neurons in these regions. That the dorsal premotor, supplementary motor, and cerebellum had this property extended the original animal studies by providing candidate areas that could also contain mirror neurons (Keysers & Gazzola, 2009).

However, fMRI suffers from an important caveat. That a voxel is active both during action observation and action execution is compatible with the voxel containing mirror neurons, but is not a proof of the presence of mirror neurons in this voxel: a voxel contains millions of neurons and can be active in both conditions despite some of its neurons only responding during observation and others only during execution and none

responding in both conditions (Gazzola & Keysers, 2009; Zaki, Wager, Singer, Keysers, & Gazzola, 2016). It thus remains essential for single cell recordings to explore the presence of mirror neurons in these regions. This has been done for the dorsal premotor cortex (Cisek & Kalaska, 2004) and the supplementary motor cortex (Mukamel, Ekstrom, Kaplan, Iacoboni, & Fried, 2010), confirming in both cases the presence of mirror neurons.

In addition to these regions involved in motor programming, voxels in the somatosensory system also show shared activations, including voxels in the posterior sector of the primary somatosensory cortex (Brodmann Area 1 and 2) and in the secondary somatosensory cortex (Gazzola & Keysers, 2009). Importantly, the pattern of activity across voxels in these regions contains specific information about what action someone else is performing, as revealed by multivoxel pattern classification (Etzel, Gazzola, & Keysers, 2008). This triggered the realization that the brain of the witness may not only trigger motor programs that would enable not only the performance of similar actions, but also somatosensory representations of what it would feel like to move in the observed way (Keysers, Kaas, & Gazzola, 2010). The activation of somatosensory cortices while witnessing the actions of others has been confirmed in a number of fMRI studies (Caspers, Zilles, Laird, & Eickhoff, 2010).

In addition, studies have also shown that sensing touch on one's own skin and witnessing someone else being touched triggers brain activity in overlapping voxels in the secondary somatosensory cortex (Keysers et al., 2004), adding evidence for the involvement of the somatosensory system in processing the somatosensory experiences of others.

Jointly, this data provides a potential neural correlate for kinesthetic empathy: when we witness the movements of others, our brain creates a representation of that movement by triggering representations of the motor programs and somatosensory feedback we would normally experience when performing similar movements or experiencing similar tactile input. The synaptic wiring necessary to map the sight of such movements to their somatosensory and motor representations could be acquired via Hebbian learning (Del Giudice, Manera, & Keysers, 2009; Keysers and Gazzola, 2014b; Keysers & Perrett, 2004; Keysers, Perrett, & Gazzola, 2014).

In a recent experiment, we further explored if these somatosensory and motor representations are functionally segregated in the brain, or strongly interact during action observation. We perturbed activity in the somatosensory cortex using transcranial magnetic stimulation (TMS) while measuring responses in the rest of the brain. We found that TMS-induced changes in brain activity in the somatosensory cortex caused a change in brain activity in the motor cortices as well (including the ventral and dorsal premotor cortex), showing that rather than generating separate motor and somatosensory representations of other people's actions, the brain seems to create an integrated somatosensory–motor representation

(Valchev, Gazzola, Avenanti, & Keysers, 2016). This is a striking match to the concept of kinesthetic empathy put forward by philosopher Theodor Lipps to explain how we enjoy art by feeling ourselves into the object of our contemplation, performing an inner mimesis, feeling ourselves moving as if we were the character depicted in the art, and thereby literally moved (Lipps, 1903).

Viewing facial expressions also triggers pattern of activity in somatosensory and motor structures that resemble the pattern when producing similar facial expressions (van der Gaag, Minderaa, & Keysers, 2007). In addition, viewing bodily expressions of emotions triggers activity in the motor system of the observer (Borgomaneri, Gazzola, & Avenanti, 2012), including activations in the premotor and parietal regions involved in moving the body (de Gelder et al., 2010). Accordingly, these motor structures could contain representations of the actions of others that could play a key role in helping us decode what other people are feeling (Jabbi & Keysers, 2008; Keysers, 2011).

THE NEURAL CORRELATES OF EMOTIONAL EMPATHY

Much of what we know about how our brain allows us to share the emotions of others takes its route from a number of fMRI studies in the early 2000s that employed an experimental design similar to those that had mapped the putative human mirror neuron system and employing a simple rationale: If empathy involves triggering a particular emotion in the witness because of witnessing that emotion in an object, a brain region involved in this ability should be activated both when the participant experiences a specific emotion himself and when witnessing the same emotion experienced by another. The first published experiment to use this logic induced the emotion of disgust in participants using unpleasant odors and then let the same participants witness the disgusted facial expressions of others. Voxels involved in experiencing and witnessing disgust were then identified separately, and the two maps were overlaid. Two brain regions had voxels that were common to witnessing and experiencing disgust (Wicker et al., 2003). The first was the anterior insula (AI) at the transition with the inferior frontal gyrus. The second was the rostral cingulate cortex (rCC) at the border between what is often referred to as the mid and anterior cingulate. The second experiment had participants experience a mild electric shock to induce pain in the self, and showed the same participants a visual cue that indicated that their romantic partner, who was also in the scanner room, received an electroshock (Singer et al., 2004). Again, the AI and rCC had voxels activated in both cases. Numerous studies have since confirmed the vicarious activation of the

AI and rCC during witnessing the emotions of others, be it for disgust (Jabbi, Bastiaansen, & Keysers, 2008; Jabbi & Keysers, 2008; Jabbi, Swart, & Keysers, 2007) or pain (Lamm, Decety, & Singer, 2011).

A key observation of that literature is that participants that report experiencing more empathy in their lives, as captured by pen-and-paper self-report questionnaires, show stronger activation of the AI and rCC when witnessing the disgust (Jabbi et al., 2007) and pain (Engen & Singer, 2013; Lamm et al., 2011; Singer et al., 2004) of others. This correlation between AI and rCC activity and empathy scores has been used as a key argument to suggest that the AI and rCC activations might be the neural basis for the empathy reported by participants. In addition, psychopaths, who are known for reduced empathy, also show reduced activation levels in these structures while perceiving the pain of others (Meffert, Gazzola, den Boer, Bartels, & Keysers, 2013). Whether the same is true for autism remains less clear (Bird et al., 2010; Hadjikhani et al., 2014).

The AI and rCC have been found to be involved in the experience and witnessing of aversive emotions such as disgust and pain, independently of how these emotions are witnessed and experienced, leading to considering them the core component of affect sharing (Engen & Singer, 2013; Lamm et al., 2011). Based on how witnesses infer the emotions of others, these regions receive their functional input through different routes. When the emotion is witnessed via facial expressions, as mentioned earlier, ventral premotor regions in the inferior frontal gyrus and somatosensory cortices that overlap with the regions involved in generating similar facial expressions are activated (van der Gaag et al., 2007), and these regions seem to provide the main functional connectivity to the AI (Jabbi & Keysers, 2008). If the states of others are inferred via linguistic narratives, these structures are embedded into a network of language-related areas that are likely to trigger the AI and rCC activation (Jabbi et al., 2008; Lerner, Honey, Silbert, & Hasson, 2011). Similarly, for pain, when pain is perceived from facial expressions or images of body parts in pain, the premotor regions of the inferior frontal gyrus are also activated (Lamm et al., 2011). In addition, if the somatic causes of pain are in the focus of the stimuli, somatosensory brain regions, including the SI and SII, are additionally activated (Keysers et al., 2010).

A much smaller number of studies have explored the neural basis of empathy for positive emotions, including joy and rewards. We have shown that experiencing and observing others experience pleasant tastes triggers activations in the gustatory cortex in the AI (Jabbi et al., 2007). We have also shown that witnessing others receiving the feedback that they had chosen a correct action overlaps with brain regions responding to similar positive outcomes for the participant's own actions, in particular in a frontostriatal network known to encore reward prediction errors, suggesting that we can learn vicariously from the actions of others

by empathizing with reward prediction errors (Monfardini et al., 2013). Similarly, in a recent study, the ventral striatum was found to encode reward prediction errors independently of whether the reward was received by the self of by someone else (Lockwood, Apps, Valton, Viding, & Roiser, 2016). Hence, while the brain regions for sharing positive and negative valences differ, the brain seems equipped with systems that could endow it with the ability to share both positive and negative emotions with others.

FACTORS MODULATING AI AND rCC ACTIVATIONS IN HUMANS

Initially, empathy was studied in neuroscience as if it were a relatively automatic process, in which participants witness the emotions of others and trigger neural representations of their own emotions. This view is in line with the literature on mirror neurons in monkeys, in which activations of the monkey's own actions occur reliably within ~100 ms each time after witnessing the action of others (Keysers et al., 2003) and that had traditionally been explained in a simple, hard-wired feed-forward stream of synaptic connections (Nelissen et al., 2011). In this view, the main determinants of variance in activations are the emotions expressed in the stimuli and individual differences in empathy that would correspond to differences in the strength of the connections between sensory cortices and the AI and rCC.

More recently, a different view has emerged, which considers empathy a motivated process that participants flexibly deploy based on their self-interest (Keysers & Gazzola, 2014a,b; Zaki, 2014). In this view, in the absence of any particular motivation to empathize, participants do recruit the AI and rCC when they witness the emotions of others to a certain degree. Individuals vary in how strongly they display such spontaneous activations. However, a number of factors can dramatically down- or up-regulate these spontaneous activations (Engen & Singer, 2013; Keysers & Gazzola, 2014a,b), five of which we will review here.

Fairness: the first such factor was observed by Singer et al. (2006). She had participants play a prisoner's dilemma game with two confederates prior to scanning, with one of the confederates being fair and trustworthy and the other unfair. Thereafter, she measured AI and rCC activation while participants witnessed either of them receiving shocks. She found that male participants show much lower activations when witnessing the unfair player receiving shocks, suggesting that perceived fairness can alter AI and rCC responses to witnessed pain.

In-/Outgroup: a number of studies have shown increased empathy-related activations in response to ingroup members receiving shocks com-

pared to outgroup members. This is true if the distinction is suggested by the object having a skin color that was similar or different from that of the witness (Azevedo et al., 2013) or by the object declaring to like the same or a different football team as the witness (Hein, Silani, Preuschoff, Batson, & Singer, 2010).

Responsibility: while in most paradigms participants witness the object experience emotions that have nothing to do with the witness' own actions, more recently, paradigms have started to explore the impact of creating causal links between the witness' actions and the object's emotions. For instance Cui, Abdelgabar, Keysers, & Gazzola, 2015 created a paradigm in which both the witness and the object had to report the central character in a flanker task. The witness then saw whether he and the object performed the task correctly or not. If either of them erred, the object received an electroshock, and the witness saw the object's facial reaction via a CCTV connection. This creates situations in which the witnessed pain is (1) entirely the responsibility of the object (in cases in which the witness had performed the task correctly, but the object erred), (2) the shared responsibility of both (when both erred), or (3) entirely the witness' responsibility (when only the witness had erred). Activations in the AI and rCC were much stronger when the observed pain was entirely the responsibility of the witness compared to both other situations.

Voluntary empathy: a small number of paradigms have asked participants to deliberately up- or downregulate their empathy for the pain of others. In one of them, healthy participants and convicted psychopathic criminals first passively viewed videos of hands hurting each others and later were asked to view these video's again, but this time trying to deliberately empathize with the victims in the videos. During the initial spontaneous viewing, psychopathic individuals activated their AI, rCC, and somatosensory cortices less than the controls. When both groups were instructed to empathize, this difference disappeared (Meffert et al., 2013). This suggests that individual differences in the AI and rCC activation are not as stable as the concept of trait differences in empathy might suggest. Rather, both groups have a similar *ability* to empathize when they want to, but differ in the degree to which this ability is spontaneously deployed in a situation in which there is no incentive to empathize (Keysers & Gazzola, 2014a,b).

Prior experience: Substantial literature has shown in the domain of kinesthetic empathy that prior experience with a particular action increases activation of premotor, somatosensory, and parietal regions associated with empathizing with that action. For instance women ballet dancers show more activation for the sight of female movements, and men for the sight of male movements for which they have prior motor experience (Calvo-Merino, Grèzes, Glaser, Passingham, & Haggard, 2006). Similarly,

experiencing a pressure pain prior to viewing others experience such pain alters the level of the AI and rCC activation in those witnessing that pain (Preis, Schmidt-Samoa, Dechent, & Kroener-Herwig, 2013). Again, it has been argued that Hebbian learning might be the key to associating one's own emotional states with the sight of the behavioral cues that signal such states in self and other (Del Giudice et al., 2009; Keysers & Gazzola, 2014b; Keysers & Perrett, 2004).

Together this draws a picture in which participants are endowed with a certain ability to activate the AI and rCC that can be shaped by a participant's past history of emotions, but the degree to which this ability is deployed while perceiving the emotions of others is highly flexible. From an evolutionary perspective, it is important to note that emotional contagion, empathy, and sympathy (which of these is most intimately associated with the AI and rCC activations remains unclear) are thought to be associated with costs and benefits. Empathizing with the sufferance of others by itself is aversive, and sympathy will motivate the witness to help the victim, which is typically costly. On the other hand, showing signs of empathy leads to increased liking, which may be beneficial if the target of empathy might later be in a position to help the witness. In addition, helping an ingroup member with whom one is likely to interact repeatedly may encourage reciprocity in the future. Finally, helping an offspring or kin has fitness benefits in a classic kin-selection sense. Accordingly, evolution should build mechanisms that deploy these phenomena in a motivated, context-dependent fashion to cash in on their benefits when such benefits prevail, but forfeit their costs if costs prevail. This would fit with empathy reductions for outgroup and unfair members, who are less likely to reciprocate prosocial actions, and with the ability to deploy empathy voluntarily. Finally, empathy can also play a key role in helping individuals find out which of their actions are beneficial vs. detrimental to others, which is key for individual's ability to function in groups. If empathy is seen as such a learning signal that punishes vicariously behaviors that harm others and rewards vicariously behaviors that benefit others, this signal should be strongest when the witnessed emotion is perceived as resulting of ones own actions, which could explain the modulation by responsibility and when they affect people we care about, which would explain the ingroup effects. Overall, such a functional view of empathy also sheds light on individual differences in propensity for empathy, as observed for instance when comparing psychopathic criminals with healthy volunteers, by suggesting that some individuals may be tuned toward a more exploitational lifestyle, in which empathy should not be deployed so readily, while others may be tuned to a more cooperative lifestyle, in which empathy should be deployed more readily (Keysers & Gazzola, 2014a; Meffert et al., 2013).

LIMITATION OF OUR UNDERSTANDING OF THE NEURAL BASIS OF EMPATHY FROM fMRI AND THE NEED FOR ANIMAL STUDIES

Since 2003, hundreds of studies have investigated the neural correlates of empathy in the emotional domain using fMRI, and even just the original two studies (Singer et al., 2004; Wicker et al., 2003) have attracted over 2500 citations (based on WoS in 2016). The success of these studies lies in the fact that they suggest (1) that emotional empathy may find its neural correlate in mirror-like neurons that fire when we experience an emotion and when we witness that emotion in others, and (2) that their activity (while we witness the emotions of others) confers us with the ability to feel what they feel and motivates us to help them. However, both of these core hypotheses remain untested. This is because these experiments are conducted in humans, in which it is, as we will review below, difficult to test either of them.

It cannot be systematically tested in humans that the same neurons in the AI and rCC are active during the witnessing and experiencing of emotions. A voxel if active both during the witnessing and experiencing of a particular emotion is compatible not only with the same neurons within it triggering activation in both cases, but also with different neurons in the same voxel recruited in both cases. If the latter were the case, the brain would retain a separation of the neural basis of our own emotions and those of others, which would thus fail to resemble what happens in the motor system, and thus fail to provide a convincing correlate of empathy. Although surgical procedures for the treatment of epilepsy offer occasional opportunities to record from single neurons in these regions in humans (Hutchison et al., 1999), they seldom offer the opportunity to carefully characterize the response pattern of neurons over multiple hours of testing, making it difficult to ascertain that neurons truly selectively represent a particular emotional experience (e.g., pain) during the observation and experience of an emotion. If brief testing then reveals the presence of some neurons recruited by both the experience and observation of emotions, we are left wondering what it is that these neurons make us share with others (Zaki et al., 2016). Indeed, if a neuron were to respond to pain when experienced by the self and the other, one is tempted to see it as potentially endowing us with the ability to share pain with others. However, if the same neuron responds when we experience disgust or fear, it would no longer allow us to specifically share pain with others, but rather a more general negative state of aversion. If it were to also respond to the observation and experience of a salient pleasant state (e.g., winning 100$), it would merely allow us to share arousal with others. Fine-grained characterizations of the tuning curve (i.e., to what a neuron does and does not respond) during both the

observation and experience of emotions is thus key to understanding the representational mechanisms that underpin our ability to share and understand the emotions of others, but requires time scales of testing that are typically unavailable in humans.

The second assumption, namely that it is the activations of the AI and rCC while witnessing the emotions of others that allow us to share their emotions and motivates our social behavior, is equally untested because these regions are relatively deep structures, over 3 cm from the surface of the cortex. Although some rare patients exist with lesions in the AI that also show reductions in empathy and the ability to recognize disgust (Adolphs, Tranel, & Damasio, 2003; Calder, Keane, Manes, Antoun, & Young, 2000), such lesions are not restricted to a particular brain region and induce compensatory cortical plasticity that make it difficult to associate the activity of a particular region to a specific psychological function. To test the causal contribution of these regions precisely, one would need to modulate their activity focally and show that this would modulate our ability to share the emotions of others. This would be essential as well to determine whether the activation of particular brain regions might be associated specifically with particular forms of empathy—would it trigger a mere emotional contagion, empathy proper, or sympathy? This is difficult to do, because the state-of-the-art methods for noninvasive brain manipulation in humans, including TMS or transcranial direct current stimulation (tDCS), are currently unable to modulate brain activity so deep without having larger effects on the cortices that are closer to the surface.

Finally, we have observed over the last decade a multitude of individual differences in empathy-like phenomena across individuals and patient groups. What causes underlie these differences? What synaptic plasticity mechanisms may differ across individuals to account for differences in the association between the perception of their states and the emotions we feel as a result? What genes associated with these differences truly cause these differences?

Accordingly, the past decade of human neuroscience of empathy has played a key role in developing strong hypotheses about the brain regions and mechanisms that might be involved in generating our subjective sense of what other people do and feel. This work has also pointed out a number of situational factors that modulate these brain activations. Testing these hypotheses, however, now urgently requires us to develop animal models of empathy in which we can leverage the powerful techniques of animal neuroscience to image (e.g., using genetically encoded calcium indicators such as GCamP6), record (e.g., using silicon probes in freely moving animals), and modulate (e.g., using chemical and optogenetic activations and deactivations) the cellular basis of these phenomenon and generate a truly mechanistic understanding of these phenomena.

FUNDING

Christian Keysers is supported by the European Research Council of the European Commission (ERC-StG-312511) and the Netherlands Organization for Scientific Research (VICI grant 453-15-009). Valeria Gazzola is supported by the Netherlands Organization for Scientific Research (VIDI grant 452-14-015), and the Brain and Behavior Research Foundation (NARSAD young investigator grant 22453).

References

Adolphs, R., Tranel, D., & Damasio, A. R. (2003). Dissociable neural systems for recognizing emotions. *Brain and Cognition*, *52*(1), 61–69. doi: 10.1016/S0278-2626(03)00009-5.

Atsak, P., Orre, M., Bakker, P., Cerliani, L., Roozendaal, B., Gazzola, V., & Keysers, C. (2011). Experience modulates vicarious freezing in rats: A model for empathy. *PLoS One*, *6*(7), e21855. doi: 10.1371/journal.pone.0021855.

Azevedo, R. T., Macaluso, E., Avenanti, A., Santangelo, V., Cazzato, V., & Aglioti, S. M. (2013). Their pain is not our pain: Brain and autonomic correlates of empathic resonance with the pain of same and different race individuals. *Human Brain Mapping*, *34*(12), 3168–3181. doi: 10.1002/hbm.22133.

Ben-Ami Bartal, I., Decety, J., & Mason, P. (2011). Empathy and pro-social behavior in rats. *Science (New York, NY)*, *334*(6061), 1427–1430. doi: 10.1126/science.1210789.

Bird, G., Silani, G., Brindley, R., White, S., Frith, U., & Singer, T. (2010). Empathic brain responses in insula are modulated by levels of alexithymia but not autism. *Brain*, *133*(5), 1515–1525. doi: 10.1093/brain/awq060.

Borgomaneri, S., Gazzola, V., & Avenanti, A. (2012). Motor mapping of implied actions during perception of emotional body language. *Brain Stimulation*, *5*(2), 70–76. doi: 10.1016/j.brs.2012.03.011.

Calder, A. J., Keane, J., Manes, F., Antoun, N., & Young, A. W. (2000). Impaired recognition and experience of disgust following brain injury. *Nature Neuroscience*, *3*(11), 1077–1078. doi: 10.1038/80586.

Calvo-Merino, B., Grèzes, J., Glaser, D. E., Passingham, R. E., & Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Current Biology: CB*, *16*(19), 1905–1910. doi: 10.1016/j.cub.2006.07.065.

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*(3), 1148–1167. doi: 10.1016/j.neuroimage.2009.12.112.

Cisek, P., & Kalaska, J. F. (2004). Neural correlates of mental rehearsal in dorsal premotor cortex. *Nature*, *431*(7011), 993–996. doi: 10.1038/nature03005.

Cui, F., Abdelgabar, A. R., Keysers, C., & Gazzola, V. (2015). Responsibility modulates pain-matrix activation elicited by the expressions of others in pain. *NeuroImage*, *114*, 371–378. doi: 10.1016/j.neuroimage.2015.03.034.

Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews*, *3*(2), 71–100. doi: 10.1177/1534582304267187.

Del Giudice, M., Manera, V., & Keysers, C. (2009). Programmed to learn? The ontogeny of mirror neurons. *Developmental Science*, *12*(2), 350–363. doi: 10.1111/j.1467-7687.2008.00783.x.

Engen, H. G., & Singer, T. (2013). Empathy circuits. *Current Opinion in Neurobiology*, *23*(2), 275–282. doi: 10.1016/j.conb.2012.11.003.

Etzel, J. A., Gazzola, V., & Keysers, C. (2008). Testing simulation theory with cross-modal multivariate classification of fMRI data. *PLoS One*, *3*(11), e3690. doi: 10.1371/journal.pone.0003690.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: From action organization to intention understanding. *Science (New York, NY)*, *308*(5722), 662–667. doi: 10.1126/science.1106138.

van der Gaag, C., Minderaa, R. B., & Keysers, C. (2007). Facial expressions: What the mirror neuron system can and cannot tell us. *Social Neuroscience*, *2*(3–4), 179–222. doi: 10.1080/17470910701376878.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain: A Journal of Neurology*, *119*(Pt 2), 593–609. doi: 10.1093/brain/119.2.593.

Gazzola, V., & Keysers, C. (2009). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: Single-subject analyses of unsmoothed fMRI data. *Cerebral Cortex*, *19*(6), 1239–1255. doi: 10.1093/cercor/bhn181.

de Gelder, B., Van den Stock, J., Meeren, H. K. M., Sinke, C. B. A., Kret, M. E., & Tamietto, M. (2010). Standing up for the body Recent progress in uncovering the networks involved in the perception of bodies and bodily expressions. *Neuroscience & Biobehavioral Reviews*, *34*(4), 513–527. doi: 10.1016/j.neubiorev.2009.10.008.

Graziano, M. S. A. (2016). Ethological action maps: A paradigm shift for the motor cortex. *Trends in Cognitive Sciences*doi: 10.1016/j.tics.2015.10.008.

Greene, J. T. (1969). Altruistic behavior in the albino rat. *Psychonomic Science*, *14*(1), 47–48. doi: 10.3758/BF03336420.

Hadjikhani, N., Zürcher, N. R., Rogier, O., Hippolyte, L., Lemonnier, E., Ruest, T., & Prkachin, K. M. (2014). Emotional contagion for pain is intact in autism spectrum disorders. *Translational Psychiatry*, *4*(November 2013), e343. doi: 10.1038/tp.2013.113.

Hein, G., Silani, G., Preuschoff, K., Batson, C. D., & Singer, T. (2010). Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. *Neuron*, *68*(1), 149–160 10.1016/j.neuron.2010.09.003.

Hutchison, W. D., Davis, K. D., Lozano, A. M., Tasker, R. R., & Dostrovsky, J. O. (1999). Pain-related neurons in the human cingulate cortex. *Nature Neuroscience*, *2*, 403–405.

Jabbi, M., Bastiaansen, J., & Keysers, C. (2008). A Common anterior insula representation of disgust observation, experience and imagination shows divergent functional connectivity pathways. *PLoS One*, *3*(8), e2939. doi: 10.1371/journal.pone.0002939.

Jabbi, M., & Keysers, C. (2008). Inferior frontal gyrus activity triggers anterior insula response to emotional facial expressions. *Emotion (Washington, DC)*, *8*(6), 775–780. doi: 10.1037/a0014194.

Jabbi, M., Swart, M., & Keysers, C. (2007). Empathy for positive and negative emotions in the gustatory cortex. *NeuroImage*, *34*(4), 1744–1753. doi: 10.1016/j.neuroimage.2006.10.032.

Jola, C., Ehrenberg, S., & Reynolds, D. (2012). The experience of watching dance: Phenomenological–neuroscience duets. *Phenomenology and the Cognitive Sciences*, *11*(1), 17–37. doi: 10.1007/s11097-010-9191-x.

Keysers, C. (2011). *The Empathic Brain*. Amsterdam, The Netherlands: Social Brain Press.

Keysers, C., & Gazzola, V. (2007). Integrating simulation and theory of mind: From self to social cognition. *Trends in Cognitive Sciences*, *11*(5), 194–196. doi: 10.1016/j.tics.2007.02.002.

Keysers, C., & Gazzola, V. (2009). Expanding the mirror: Vicarious activity for actions, emotions, and sensations. *Current Opinion in Neurobiology*, *19*(6), 666–671. doi: 10.1016/j.conb.2009.10.006.

Keysers, C., & Gazzola, V. (2014a). Dissociating the ability and propensity for empathy. *Trends in Cognitive Sciences*, *18*(4), 163–166. doi: 10.1016/j.tics.2013.12.011.

Keysers, C., & Gazzola, V. (2014b). Hebbian learning and predictive mirror neurons for actions, sensations and emotions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1644), 20130175. doi: 10.1098/rstb.2013.0175.

Keysers, C., Kaas, J. H., & Gazzola, V. (2010). Somatosensation in social perception. *Nature Reviews. Neuroscience*, *11*(6), 417–428. doi: 10.1038/nrn2833.

Keysers, C., Kohler, E., Umiltà, M. A., Nanetti, L., Fogassi, L., & Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research*, *153*(4), 628–636. doi: 10.1007/s00221-003-1603-5.

Keysers, C., & Perrett, D. I. (2004). Demystifying social cognition: A Hebbian perspective. *Trends in Cognitive Sciences*, *8*(11), 501–507. doi: 10.1016/j.tics.2004.09.005.

Keysers, C., Perrett, D. I., & Gazzola, V. (2014). Hebbian learning is about contingency, not contiguity, and explains the emergence of predictive mirror neurons. *Behavioral and Brain Sciences*, *37*(2), 205–206. doi: 10.1017/S0140525X13002343.

Keysers, C., Wicker, B., Gazzola, V., Anton, J. -L., Fogassi, L., & Gallese, V. (2004). A touching sight: SII/PV activation during the observation and experience of touch. *Neuron*, *42*(2), 335–346 Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15091347.

Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: Action representation in mirror neurons. *Science (New York, NY)*, *297*(5582), 846–848. doi: 10.1126/science.1070311.

Lamm, C., Decety, J., & Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *NeuroImage*, *54*(3), 2492–2502. doi: 10.1016/j.neuroimage.2010.10.014.

Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(8), 2906–2915. doi: 10.1523/JNEUROSCI.3684-10.2011.

Lipps, T. (1903). *Asthetik: Psychologie des Schönen und der Kunst*. Hamburg: Leopold Voss.

Lockwood, P. L., Apps, M. A. J., Valton, V., Viding, E., & Roiser, J. P. (2016). Neurocomputational mechanisms of prosocial learning and links to empathy. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(35), 9763–9768. doi: 10.1073/pnas.1603198113.

Meffert, H., Gazzola, V., den Boer, J. A., Bartels, A. A. J., & Keysers, C. (2013). Reduced spontaneous but relatively normal deliberate vicarious representations in psychopathy. *Brain: A Journal of Neurology*, *136*(Pt 8), 2550–2562. doi: 10.1093/brain/awt190.

Monfardini, E., Gazzola, V., Boussaoud, D., Brovelli, A., Keysers, C., & Wicker, B. (2013). Vicarious neural processing of outcomes during observational learning. *PLoS One*, *8*(9), e73879. doi: 10.1371/journal.pone.0073879.

Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current Biology*, *20*(8), 750–756. doi: 10.1016/j.cub.2010.02.045.

Nelissen, K., Borra, E., Gerbella, M., Rozzi, S., Luppino, G., Vanduffel, W., & Orban, G. A. (2011). Action observation circuits in the macaque monkey cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(10), 3743–3756. doi: 10.1523/JNEUROSCI.4803-10.2011.

Panksepp, J., & Panksepp, J. B. (2013). Toward a cross-species understanding of empathy. *Trends in Neurosciences*doi: 10.1016/j.tins.2013.04.009.

Preis, M. A., Schmidt-Samoa, C., Dechent, P., & Kroener-Herwig, B. (2013). The effects of prior pain experience on neural correlates of empathy for pain: An fMRI study. *Pain*, *154*(3), 411–418. doi: 10.1016/j.pain.2012.11.014.

Rice, G. E., & Gainer, P. (1962). "Altruism" in the albino rat. *Journal of Comparative and Physiological Psychology*, *55*, 123–125 Retrieved from http://www.ncbi.nlm.nihgov/pubmed/14491896.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science (New York, NY)*, *303*(5661), 1157–1162 10.1126/science.1093535.

Singer, T., Seymour, B., O'Doherty, J. P., Stephan, K. E., Dolan, R. J., & Frith, C. D. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature*, *439*, 466–469.

Valchev, N., Gazzola, V., Avenanti, A., & Keysers, C. (2016). Primary somatosensory contribution to action observation brain activity-combining fMRI and cTBS. *Social Cognitive and Affective Neuroscience*doi: 10.1093/scan/nsw029.

de Vignemont, F., & Singer, T. (2006). The empathic brain: How, when and why? *Trends in Cognitive Sciences*, *10*(10), 435–441. doi: 10.1016/j.tics.2006.08.008.

Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in My insula: The common neural basis of seeing and feeling disgust. *Neuron*, *40*(3), 655–664 Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14642287.

Wispé, L. (1986). The distinction between sympathy and empathy: To call forth a concept, a word is needed. *Journal of Personality and Social Psychology*, *50*(2), 314–321. doi: 10.1037/0022-3514.50.2.314.

Zaki, J. (2014). Empathy: A motivated account. *Psychological Bulletin*, *140*(6), 1608–1647. doi: 10.1037/a0037679.

Zaki, J., Wager, T. D., Singer, T., Keysers, C., & Gazzola, V. (2016). The anatomy of suffering: Understanding the relationship between nociceptive and empathic pain. *Trends in Cognitive Sciences*, *20*(4), 249–259. doi: 10.1016/j.tics.2016.02.003.