Review article

Imaging empathy and prosocial emotions

Claus Lamm *, Markus Rütgen, Isabella C. Wagner

Social, Cognitive and Affective Neuroscience Unit, Department of Basic Psychological Research and Research Methods, Faculty of Psychology, University of Vienna, Liebigasse 5, 1010 Vienna, Austria

HIGHLIGHTS

- Empathy is a multi-faceted construct.
- Affect sharing, self-other distinction, compassion, prosocial behavior are distinct phenomena.
- These phenomena differ with respect to their neural mechanisms.
- Progress in their understanding requires an integrative approach.

ARTICLE INFO

Article history:
Received 10 February 2017
Received in revised form 17 May 2017
Accepted 28 June 2017
Available online 29 June 2017

Keywords:
Empathy
Compassion
Self-other distinction
Prosocial behavior
fMRI
Neuroscience

ABSTRACT

Empathy is a multi-faceted construct with important implications for social behavior. Based on a selective review of the neuroscientific evidence collected in humans, the present paper discusses the neural representations underlying affect sharing, its relation to mentalizing, the importance of self-other distinction, the distinction between empathy, sympathy and compassion, and how these phenomena are linked to prosocial behavior. Apart from reviewing the literature, we also highlight open questions and how they might be addressed by a research approach that tries to integrate across these diverse constructs.

© 2017 Elsevier B.V. All rights reserved.

Contents

1. Empathy – a multi-faceted construct .......................................................... 49
2. Affect sharing and “shared representations” .................................................. 50
3. Self-other distinction ...................................................................................... 51
4. Sympathy and compassion ............................................................................. 51
5. Prosocial behavior ......................................................................................... 52
6. Conclusion ...................................................................................................... 52
Acknowledgments ............................................................................................ 52
References ........................................................................................................ 52

Recent years have brought a surge of studies on the neural underpinnings of empathy and on how this social skill motivates us (inter)act in a prosocial manner. The present review aims to provide a selective overview of this research, and of remaining open questions. Before taking that journey, it is important to note that empathy is a multi-faceted and complex construct requiring disambiguation from various related terms.

1. Empathy – a multi-faceted construct

While many different definitions of empathy exist [1][1 for a review], the one proposed in [2] highlights several aspects on which we will focus on in this paper, which are affect sharing, mentalizing, self-other distinction, sympathy and compassion, and prosocial behavior. More specifically, according to this definition, empa-
A great deal of neuroscientific research on empathy has focused on affect sharing. One influential view, the so-called “shared representations account of empathy”, suggests that empathy for a certain emotion engages neural processes that also underpin the first-hand experience of that emotion [6,7]. This account was initially fueled by the robust finding that empathy for pain activates mid-cingulate and anterior insular cortices – i.e., areas that are also activated when pain is experienced directly [8 for meta-analysis,9,10; see also Fig. 1]. Later studies revealed similar findings, also for positive affect [6,11,12]. Shared activations have often been interpreted in the sense of shared representations, which would imply that similar or equivalent neuro-cognitive processes are underpinning affect sharing. However, inferring representations from neural activation is by no means straightforward, in particular due to the inherent limitations of the so-called mass-univariate analyses of fMRI data which are standard in the field [13–15]. One approach that has thus gained particular momentum is multi-voxel pattern analysis (MVPA). It provides information about underlying neural representations by investigating patterns of activation rather than mass-univariate activation levels [16,17]. Results from studies that used this method for detecting shared or dissociable representations during affect sharing are somewhat complex, though. While some studies presented evidence for shared representations, meaning that activation patterns in response to empathy for pain could be predicted (or “cross-classified”) based on patterns in response to the first-hand experience of pain [18,19], others who used a different analysis approach, suggested a lack of shared representations between the two conditions [15, see also 20]. Moreover, the latter study also included additional analyses of the cross-classification accuracy, concluding that classification might be driven by domain-general affect rather than pain specifically [see also 21].

Of note are also the discrepancies between mass-univariate and MVPA results, with e.g. the former in [15] suggesting activation of areas related to empathy, and the latter indicating engagement of areas classically related to mentalizing. This suggests that the different analyses approaches are sensitive to different kinds of information within the fMRI data. In the same vein as mass-univariate analyses, MVPA results thus need to be interpreted with some caution, as our understanding of the exact neural processes reflected by both univariate activation measures and voxel patterns is still limited [see 22, and responses for an excellent discussion].

Neither mass-univariate nor MVPA approaches provide a mechanistic explanation of the role of shared or dissociable representations during empathy. A further suggestion is thus to combine fMRI, which is correlational by nature, with research methods that allow for causal interpretations. One such method are lesions studies, which suggest that patients with lesions of anterior insular cortex exhibit decreased empathic skills [23]. While this finding points towards a necessary role of this brain region in affect sharing, it also suffers from inherent limitations of the lesion approach, such as that lesions usually encompass also bordering regions, or that conclusions are limited to the specific patient population [24, for review]. Another possibility to causally investigate shared or dissociable representations of pain or empathy for pain is to induce a so-called “placebo empathy analgesia” in healthy partic-

Fig. 1. Schematic overview of brain regions associated with the central concepts discussed in this review article. An integration and synthesis of how these regions actually interact in generating prosocial emotions and behavior is, however, still missing (see text for major open research questions). Central concepts and their definitions (see [11] and [4], for review): affect sharing: isomorphic sharing of the affective state of another person; compassion, sympathy: feeling concern for another person and his or her suffering; mentalizing: ability to make inferences about others’ mental states; self-other distinction: ability to distinguish between cognitive, sensory and affective states pertaining to ourselves and to others. Red shadings mark the approximate locations of respective regions: AI, anterior insula; aMCC, anterior mid-cingulate cortex; mPFC, medial prefrontal cortex; TPJ, temporo-parietal junction; VS, ventral striatum; L, left; R, right; RH, right hemisphere.
patients. We recently used this approach to show that the intake of an alleged “painkiller” not only reduced first-hand pain, but also empathy for pain. This was accompanied by reduced activity in the shared neural networks that coded for the two experiences. Moreover, pharmacologically manipulating opioid activity by means of the opioid receptor antagonist naloxone reverted these effects, suggesting that similar neurotransmitters are engaged during pain and empathy for pain [25,26]. Similar findings on self-reports of empathy were observed following administration of the painkiller acetaminophen [27]. However, further research is needed to clarify whether this (psycho)pharmacological approach unequivocally speaks for shared representations. For instance, placebo analgesia may not only modulate pain processing, but it might have general effects on how we process and experience negative affect. While this would not contest affect sharing per se, it would question the specificity of the shared representations (i.e., whether it is pain vs. general aversive affect that is shared). Also, it needs to be shown whether the findings for pain apply also to other affective experiences and emotions. Additionally, the role of shared representations has been investigated with respect to the sharing of sensorimotor processes, with causal approaches becoming more widely used as well [28], and suggesting that sensorimotor resonance may also contribute to affect sharing [29,30 for reviews].

Finally, research is needed that assesses more clearly the implications of shared representations for empathic understanding, and the behavior motivated by it [see also 31]. While it seems reasonable to assume that sharing other’s affect based on one’s own affective representations enables us to understand others better, such an experiential understanding might also be counterproductive in some cases— as indeed shown by the placebo analgesia studies reviewed above— or other situations in which self-other distinction fails.

3. Self-other distinction

Self-other distinction is important to avoid that our feelings bias how accurately we share and understand another person’s affective state [32]. A lack of self-other distinction can have deleterious effects on prosocial behavior, as its failure may increase personal distress, a self-related aversive response detractions the focus on our partner’s suffering towards our own distress and its regulation [3]. One area with which self-other distinction has consistently been associated with is the right temporoparietal junction (rTPJ), see Fig. 1). While the precise function of this brain area still awaits clarification [14, for a review, 33], recent findings suggest that self-other distinction in the affective domain might be related to subdivisions of the rTPJ that are distinct from those engaged during self-other distinction in the motor or the cognitive domain [34,35]. More specifically, in experimental paradigms that were tailored to pinpoint how self-related affect biases empathy, [36] as well as [37] have shown that overcoming such a bias, and thus engaging in affective self-other distinction, is associated with increased activation in the right supramarginal gyrus (rSMG), a cortical area slightly anterior to the commonly reported rTPJ activation increases during self-other distinction. Moreover, applying repetitive transcranial magnetic stimulation in an inhibitory fashion above the rSMG decreased affective self-other distinction [36], and functional connectivity analyses revealed that this was associated either with 1) a putative multimodal social integration mechanism as shown by increased connectivity from the rSMG to areas processing either self- or other-related sensory information [36], or 2) with prefrontal controls that are supposedly engaged in affect inhibition [37].

However, several questions regarding the precise functions of self-other distinction and its neural implementation, await clarification. It is unclear whether the rTPJ and rSMG have different functions, or whether they have similar functions, but are anatomically segregated because of differences in the type of information that is preferentially being processed (e.g., sensory-motor, cognitive for rTPJ, or affective for rSMG). The latter interpretation is in line with resting-state functional connectivity analyses showing that the posterior and anterior rTPJ (including rSMG) display different connectivity profiles with regions that are associated with cognitive and affective processing [38], respectively. Moreover, it needs to be shown how rTPJ and rSMG engagement during self-other distinction relates to other cognitive functions, both in the social and non-social domain. For instance, the rTPJ is also engaged during attention reorienting, which might constitute a mechanism by which self-other distinction is achieved— i.e., by shifting attention from representations pertaining to the self to those of the other [39,40]. The TPJ also seems to have a prominent role during mentalizing, but this was linked to making inferences about other’s inner states rather than to keep them distinct from one’s own [35]. Finally, how exactly self-other distinction is neurally implemented remains elusive, with possibilities ranging from a simple “tagging” mechanism that allows us to keep self- and other related representations separated, to mechanisms that actively manipulate (e.g., inhibit or enhance) these representations [14, for a detailed review]. Taken together, the complexity of these findings highlights the need for a unitary framework that integrates across the many different cognitive and affective functions during which activation in the TPJ and adjacent SMG has been reported.

4. Sympathy and compassion

On a conceptual level, empathy and sympathy (or the related terms of care and compassion) should be regarded as distinct phenomena [see 4 for in-depth discussion]. The requirement of this distinction has recently become obvious again in a controversy on how empathy is related to morality, which may partially be resolved by differences in the use of the terms “empathy” and “compassion” by different scholars [13,41,42]. Based on neural data, we have been gaining an increasingly detailed understanding of what distinguishes empathy from sympathy and compassion. Early work in this respect focused on cross-sectional studies of brain function and neuroplasticity in Buddhist monks, i.e. persons engaged in the persistent cultivation of meditation practices tailored to increase compassion [e.g., 43]. More recently, evidence was presented that even short-term meditation training can have similar effects. Although paradoxically at a first glance, results showed that the brain networks engaged during empathy and compassion overlap only scarcely [44 for review, 45], as compassion engages areas that are traditionally considered to be part of the human reward system. For instance, comparing meditation training that increases either empathy (as defined above) or compassion, Klimecki and colleagues reported increased negative affect for the former and increased positive affect in the latter [46,47]. Moreover, compassion training led to higher activation in ventral striatum and medial orbito-frontal cortex. This may result from the specific meditative practice trained, which entailed the generation of feelings of warmth and a wish for well-being when imagining (the suffering of) others. Hence, it remains to be shown whether feelings of care and compassion activate similar networks also without such compassion training. This indeed seems to be suggested by a very recent study that disentangled the feelings of empathy, in the sense of affect sharing, and vicarious distress from those related to care, compassion, and prosocial behavior [48].

Since these studies suggest that empathy and compassion are distinctly implemented on the neural level this raises the question of how they are related and interact— e.g., whether they
build up on each other, or whether they can arise independently. Answering this question is highly relevant to better understand the role of individual differences in prosocial emotions and behaviors. For instance, research on psychopathy suggests that individuals affected by these disorders might not suffer from a principled lack of empathy, and certainly not of mentalizing, but rather from a lack of compassion and possibly of how shared affect is transformed into compassion [49]. Research on the effects of expertise has further shown that experienced acupuncturists do not engage brain regions associated with affect sharing when confronted with the pain of others, perhaps related to inhibitory prefrontal control and emotion regulation, as well as mentalizing [50]. This suggests that experience with aversive acts can result in a down-regulation of affect sharing, but it has not yet been investigated whether this also affects empathy and concern, or how this relates to accurate patient treatment [see also 51]. Future studies should therefore assess how individual differences in mentalizing, empathy, and compassion interact to affect social interaction and behavior, how this is affected by individual differences in personality and experience, and how neural systems underpin the behavioral outcome.

5. Prosocial behavior

What motivates prosocial behavior certainly bears the strongest potential implications that neuroscience research on social emotions might have. After all, we do not only want to know what makes us share and understand the feelings of others, but also how this motivates us to behave towards others. One particularly interesting trend in the literature is the use of computational models to formally explain behavior [52–54]. For instance, [52] modelled fMRI data during reinforcement learning to show that learning to obtain rewards for others seems to rely on similar brain structures as learning to reap rewards for oneself. The identified areas such as the subgenual anterior cingulate cortex partially overlapped with areas involved in vicarious reward [11 for review], suggesting a link between representing the positive affect of others and acting upon it for their benefit. It is also interesting that the identified areas do not include areas that are traditionally associated with empathy for negative affect, speaking for a differentiation and a specificity of affect sharing and prosocial behavior for the type, or at least the valence, of emotions [12]. A related question is how empathy, compassion, and mentalizing interact to drive prosocial behavior. For instance, one recent study demonstrated that activation in the medial prefrontal areas predicts prosocial behavior during mentalizing [55], while another investigation suggested that helping behavior is predicted by activation related to affect sharing [56]. Yet others have suggested that it seems to be sympathy and concern that drive altruistic acts [57]. In an attempt to investigate how mentalizing, empathy, sympathy, and prosocial behavior are related, we have thus recently investigated whether the effects of mentalizing on prosocial behavior are mediated by enhancing other-related mental representations rather than by self-projection mechanisms [58]. While this can be seen as a first step towards a more integrative understanding of the multiple determinants of prosocial behavior and their interaction, future studies are needed to come to more definite conclusions. These studies should also develop experimental paradigms that are more closely related to real-life prosocial behavior, e.g., through the use of virtual reality simulations [59].

6. Conclusion

Empathy is a highly complex and multi-faceted phenomenon that has fascinated scholars of various disciplines for centuries. It also attracts enormous interest by the general public, due to the implications it may have to address the major societal and political challenges we are currently facing. While having provided some major first insights, the neuroscientific investigation of this phenomenon is still in its infancy. By focusing on the many facets of empathy and its effects on social behavior, the present review aims to promote an integrative approach that attempts to connect these facets, which so far often were studied in isolation. Ultimately, this might lead to a comprehensive understanding not only of the neural mechanisms of empathy but also their biological origins, how these mechanisms develop and are shaped by experience, which interpersonal and contextual factors influence them, and how this manifests itself in moral behavior [13, 60].

Acknowledgments

The authors received financial support from the Austrian Science Fund [grant number P29150], and the Federal Ministry of Science, Research & Economy [grant name: “Interdisciplinary Translational Brain Research Cluster with Highfield MR”].

References