

Ricerche

Against a Broad Definition of “Empathy”

Sarah Songhorian

Ricevuto: 17 novembre 2014; accettato: 4 febbraio 2015

Abstract In this paper I will try to provide some arguments against a broad definition of “empathy”. Firstly, I will deal with attempts to define empathy as an umbrella concept. Then, I will try to point out the four main elements which contribute to the confusion that researchers in both the social and political as well as the scientific and philosophical domains face when dealing with empathy. In order to resolve this confusion, I suggest applying David Marr’s distinction to the field of empathy. Instead of providing an umbrella definition for empathy, which tries to account for all the data coming from different disciplines, I believe understanding that there are different levels of explanations and that different disciplines can contribute to each of them will provide a more detailed and less confused definition of empathy.

KEYWORDS: Empathy; D. Marr; de Waal; Identification; Helping-behavior; Cost/benefit Analysis.

Riassunto *Contro una definizione ampia di “empatia”* – In questo articolo cercherò di fornire alcuni argomenti contrari all’adozione di una definizione ampia di “empatia”. In primo luogo cercherò di definire la nozione di empatia come una sorta di ombrello concettuale. Proverò poi a indicare i quattro principali elementi responsabili delle confusioni di cui gli studiosi cadono vittima sia in ambito sociale e politico sia in ambito scientifico e filosofico quando si occupano di empatia. Per sgombrare il campo da questa confusione proporrò di applicare la distinzione avanzata da David Marr al campo dell’empatia. Anziché dare una definizione di empatia come una sorta di ombrello concettuale, cercando di tenere insieme tutte le evidenze provenienti da discipline diverse, sono convinta che la comprensione del fatto che possono darsi livelli differenti di spiegazione e che discipline diverse possano contribuire ognuna per proprio conto alla comprensione di questi livelli costituisca la via migliore per raggiungere una definizione di empatia più specifica e meno confusa.

PAROLE CHIAVE: Empatia; D. Marr; De Waal; Identificazione; Comportamento di aiuto; Analisi costi/benefici.



Introduction

IN RECENT TIMES, THERE HAS BEEN growing interest concerning the concept of “empathy” in the literature.¹ My aim in this paper will be to underline the utility of a conceptual clarification within this research topic.

The argument for further distinctions between different concepts and levels will come

from the critical analysis of “empathy” as a broad concept.² These accounts – aside from relevant differences – all rely on the possibility of conceiving of “empathy” as a superordinate category that includes different processes or components. Through my critical analysis, I will advocate the need for a clear distinction between related concepts. Karsten Stueber correctly noted in his overview of em-

S. Songhorian - Facoltà di Filosofia, Università San Raffaele “Vita-Salute”, via Olgettina, 58 - 20132 Milano (I). E-mail: s.songhorian@gmail.com (✉)



pathy that:

As psychologists themselves have become increasingly aware, the empirical investigation of empathy has been hindered (particularly in the beginning) by conceptual confusions and a multiplicity of definitions of the empathy concept [...] Within social psychology, this state of affairs is due to the fact that the empathy concept merged with and completely replaced the multi-dimensional concept of sympathy used by earlier psychologists and philosophers.³

I will also argue that much of the criticism that has been directed at empathy is due to this conceptual confusion and can be avoided by making a further distinction. I will use Paul Bloom’s article *Against Empathy*.⁴ as an example that illustrates this. I believe what divides Paul Bloom’s position – and that of some of his commentators – from my own is a matter of definition.

■ Criticism of Preston and de Waal’s account of “Empathy”

Let me start by outlining four main criticisms of Preston and de Waal’s account which I take to be an example of a broad view on “empathy”. In what follows, I will focus solely on those aspects of the paper that I wish to critique.

The first two criticisms deal with the authors’ intention to distinguish only between proximate and ultimate causes of behavior.⁵ In the explanation the authors give about the specific causes of empathy,⁶ it is hard to distinguish theories from data. In this respect, it is useful to consider Marr’s distinction between different levels of explanation.⁷ To consider this distinction is not to say that proximate and ultimate causes do not matter, but just to suggest that a further differentiation between the level of computational theory, the algorithmic or representational level, and that of implementation might be useful

in disentangling different aspects of empathy. As the authors themselves claim:

The authors view the term empathy broadly, similar to Hoffman, as: *any process where the attended perception of the object’s state generates a state in the subject that is more applicable to the object’s state or situation than to the subject’s own prior state or situation* [...] A process model makes empathy a superordinate category that includes all subclasses of phenomena that share the same mechanism.⁸

The subclasses of phenomena the authors have in mind are represented in their Figure 1 and include: identification, emotional contagion, “true empathy”, cognitive empathy, helping behavior, and guilt.⁹ Moreover, in a previous version of the paper, the same figure included: identification, emotional contagion, “true empathy”, cognitive empathy, perspective taking, and theory of mind.¹⁰

As Hinde correctly notes, in his commentary on Preston and de Waal’s article, the aim of integrating different phenomena in order to provide a unified explanation is definitely a relevant one, but the way in which the authors attempt to do it lacks precision and explanatory power:

Definitional clarity is not enhanced by the juxtaposition of distinct levels of discourse.¹¹

Going back to the necessity of a further distinction between levels of explanation, instances such as the following can be found in the paper:

A Perception-Action Model of empathy specifically states that *attended perception of the object’s state automatically activates the subject’s representations of the state, situation, and object, and that activation of these representations automatically primes or generates the associated automatic and somatic responses, unless inhibited*.¹²

What might be observed in instances such as the latter is the fact that the authors provide an interpretation of data coming from ethology, developmental psychology, and neuro-imaging studies that – though plausible – is not the only explanation possible. It is somehow taken for granted that certain data support a specific theoretical perspective without a deeper understanding of other possible interpretations. In the relation between theoretical accounts and data this is a common, yet dangerous, error. It might give the reader the impression that a specific set of data is the ultimate proof for a specific theory. In this respect, philosophical analysis can prove extremely useful.

The *Perception-Action Model* (PAM), as the authors themselves claim, is a theoretical framework that claims that a certain state of the object automatically activates a certain set of representations, which themselves activate, again automatically, certain somatic responses. It is clear that in the passage I just quoted the authors shift from a theoretical level to a representational one and, moreover, to an implementational one without any explanation for these moves. As I mentioned, I believe understanding that there are three different, yet connected, levels of explanation might be useful in better understanding the phenomenon of empathy. It is one thing to propose a theoretical account, such as PAM, and another to claim that in an empathic phenomenon certain representations are activated. Moreover, how these representations are implemented is yet another thing.

Obviously the concept of representation is itself particularly problematic. Moreover it is hard to understand – as Martin Hoffman noticed in his *Commentary* to Preston and de Waal's paper¹³ – how it could be the case that the process the authors describe would be useful for all of the different phenomena that they conceive as part of the concept of empathy. Namely, it is hard to figure out how an explanation that refers to representations of states, situations, and objects can be ascribed to identification and emotional contagion as

primary experiences of empathy that even infants can have. The response provided by the authors claims that their usage of “representation”:

refer[s] to the neuronal connections of the brain that store the information. As such, muscle movements, feeling states, associations, conditioning, and so on, are all mediated by representations, and they require no special cognitive abilities beyond the plasticity that exists in any central nervous system.¹⁴

Though this explanation somehow accommodates Hoffman's objection as far as basic forms of emotional sharing are concerned, it cannot accommodate the presence of different levels of explanation being displayed as a continuum. Neither does it allow for more complex phenomena for which we think that more than simply plasticity is needed to be classified as cognitive empathy.

Again, it seems difficult to provide a unified explanation that can account for all of the phenomena Preston and de Waal want to explain. Furthermore, in Preston and de Waal's work, there is not only in the problem of distinguishing between the theoretical, the representational, and the implementational levels, but there are also unjustified moves from the mechanisms we are endowed with – obviously neutral – to a specific and desirable behavior.

Thus nervous systems that respond automatically with empathy to situations where they must respond.¹⁵ The risk that this kind of assertion poses is that of a hard naturalization of moral behavior, according to which, provided there is a certain mechanism we are endowed with, it can be directly derived that a certain behavior will be displayed. Obviously, the authors do not advocate such a strong view, but there is a risk of overinterpreting the data. Both these criticisms can be avoided by means of a further distinction between levels of explanation and the difference between moral psychology – interpreted as providing us a description of the

basic mechanisms that support moral judgment and the direction of one’s behavior – and moral behavior itself.

The core idea is that the reason why we should care about the data concerning our basic capacities when doing moral philosophy is to try and identify what is morally possible for beings such as we are. In order to prescribe a behavior, we need to know that a specific kind of response to what happens to us can be displayed by the kind of individuals we are.

The other two criticisms to Preston and de Waal’s account concern specifically the idea that the better way to understand empathy and all the literature on it is to provide a “unified whole”.¹⁶ In a more recent paper, Frans de Waal¹⁷ suggests a more plausible re-interpretation of the relations between all of the phenomena that seem to be conceived as subsets of the concept of empathy in his paper with Preston. The *Russian Doll Model*¹⁸ claims that:

higher cognitive levels of empathy [are] built upon a firm, hard-wired basis, such as PAM.¹⁹

Besides the more plausible explanation that the *Russian Doll Model* entails, two main difficulties remain. On the one hand, it is assumed that the PAM represents the basic form of empathy, without any further discussion of possible alternative interpretations; on the other – more relevantly – de Waal suggest that all of the phenomena constituting the outer layers of the doll share the same core mechanisms with basic forms of empathy.

In order to show that it is not necessary for these phenomena to share the same mechanisms, let me briefly analyze the relation between “true empathy” and helping behavior. I will do this firstly by looking at the explanation Preston and de Waal propose in their article for such a link and, afterwards, by underlining the difference between the two levels of the description. In Preston and de Waal’s article, there are two possible interpretations of the connection between em-

pathy and helping behavior. On the one hand, given the examples the authors provide, it looks like the fact that rats, monkeys and infants are able to display a certain behavior constitutes evidence for their ability to empathize. In particular:

In an experiment with rhesus monkeys, subjects were trained to pull two chains that delivered different amounts of food. The experimenters then altered the situation so that pulling the chain with the larger reward caused a monkey in sight of the subject to be shocked. After the subjects witnessed the shock of the conspecific, two-thirds preferred the nonshock chain even though it resulted in half as many rewards. Of the remaining third, one stopped pulling the chains altogether for 5 days and another for 12 days after witnessing the shock of the object. These monkeys were literally starving themselves to prevent the shock to the conspecific.²⁰

The authors refer to the experiment conducted by Masserman et al.²¹ To construe empathy as a broad phenomenon does not sufficiently underline the difference between “feeling the state of another” and performing a helping or pro-social behavior because of that feeling. These are two different, though connected, phenomena. It is one thing to underscore the connection between empathy and pro-social behavior; it is quite another to consider the latter as part of the concept of the former itself.

These examples, all from empirical reports, show that individuals of many species are distressed by the distress of a conspecific and will act to terminate the object’s distress, even incurring risk to themselves.²²

The experiments are taken to be the proof of a direct and pre-reflective link between empathy and helping behavior. On the other hand, researchers try to emphasize that pro-social behavior is the result of a complex cost/benefit analysis and it is not obviously a direct and unconscious consequence of em-

pathy:

it may be more accurate to consider helping behavior as the result of a complex cost/benefit analysis on the perceived effectiveness of helping and the effect of helping on short and long-term goals.²³

It is not clear how the two definitions of the relation between empathy and helping behavior can be integrated. If the examples prove a point and they are actually evidence of empathy, then the link between empathy and helping behavior seems to be direct and it is not clear why we should concede that the latter has having to do with a cost/benefit analysis.

Obviously, the difficulty, if this version of the connection is the correct one, arises when we confront ourselves with occurrences in which we believe there is an empathic response but no helping behavior follows. Again, according to this version, the risk is that of a hard naturalization: provided we have a certain mechanism, we will act in a certain – pro-social or moral – way upon it. The question as to why we sometimes do not respond by helping others then becomes a real issue. If, on the other hand, it is true that helping behavior is a more complex phenomenon involving a cost/benefit analysis, then it is not clear why it should be counted within the scope of the concept of empathy itself.

Moreover, the authors continue to shift from one definition to the other. If the examples were not meant to be considered as evidence, despite the explicit reference to them as that kind of evidence, and the idea that helping behavior is a cost/benefit analysis was to be considered as the proper definition they wanted to put forward, it is not easy to understand how they can state the following thesis:

A process model makes empathy a superordinate category that includes all subclasses of phenomena that share the same mechanism.²⁴

If the cost/benefit analysis is conceived as

a conscious one, data from fMRI studies suggest that the mechanisms underlining it are different from those at the basis of emotional responses.²⁵

A possible solution for this apparent contradiction is that of understanding the cost/benefit analysis as an unconscious mechanism.²⁶ This interpretation might well explain basic forms of helping behavior, such as those shown by infants or primates, but it will not account for more complex and more cognitively loaded forms of helping behavior. A few examples may clarify my argument against the idea of helping behavior deriving from an unconscious cost/benefit analysis. On the one hand, there are cases in which individuals are completely other-oriented in their behavioral choices. These are obviously instances of supererogative behavior, but it is complicated to understand them as deriving from an appropriate cost/benefit analysis.

In particular, if helping behavior depends on a cost/benefit analysis, supererogative behavior cannot constitute the higher level of moral behavior, but rather a misdirected analysis that disregards risks to the first person in balancing costs and benefits. If John incurs a risk to his own life in order to save a stranger, it might well be the case that he is doing so out of altruism instead of because of a cost/benefit analysis. On the other, if the mechanism is unconscious, how can we account for occurrences of conscious cost/benefit analysis before behaving? How is it possible that for more basic forms of empathy the cost/benefit analysis is unconscious, while for those that are more complex we need to be informed of such an analysis?

We have all experienced occasions in which we were strongly considering whether to help someone or not based on the outcomes – both for ourselves and others – of such a helping behavior. So that, from experience, we can easily see that – at least sometimes – helping behavior requires a great deal of conscious analysis.

Again, the problem with these cases relates to the different levels that they account

for. From a specific behavior, we cannot straightforwardly derive which mechanism underlies it. Behavioral responses are not sufficient even to identify and distinguish different forms of emotional experience, let alone empathic occurrences: from behavioral data we cannot infer that an empathic episode is going on, there might well be many other relevant causes of the behavior.

There is also another difficulty in this connection between empathy and helping behavior: how can it account for occurrences in which we help for reasons other than emotional arousal? John can help Max because he sees him in distress and he wants to help him alleviate that feeling, or because he feels an obligation or a duty to do so, or because he believes Max will then feel an obligation to reciprocate, or because helping him will serve some other interest or goal John has. Helping behavior *per se* does not necessarily need to be linked with empathy or with an understanding of others' emotions.

The attempt to connect empathy with helping or pro-social behavior, though relevant, cannot be conducted in such a direct and straightforward way: it gives the misleading idea that we can move directly from a description of a mechanism to its behavioral consequences as if those were univocal, caused only by the former, and non-controversial. Helping behavior can be the consequence of so many different mechanisms at work that to rely simply on its connection to empathy would be reductive. I believe the more feasible way to connect helping behavior and empathy is to derive the former, at least in some relevant case, from a specific moral character to be built based on sympathy rather than on empathy as a basic mechanism. The connection between empathy and sympathy cannot be analyzed here.

This is not to say that there is no relation at all between the phenomena Preston and de Waal describe as constituting the domain of empathy – i.e. identification, emotional contagion, “true empathy”, cognitive empathy, helping behavior, and guilt – but just to sug-

gest that distinguishing between these phenomena can provide a more accurate account of all of them, of their mechanisms, and their relations. To say that there is continuity of these phenomena – as Preston and de Waal claim – does not necessarily mean that they have to be interpreted as being the same thing.

Moreover, it seems that the idea that all these phenomena fall under the same conceptual framework depends on the acceptance of a *Perception-Action Model* (PAM). So that, if PAM falls, there is no reason to believe that the rest of the theoretical framework should hold. I cannot deal with this objection here, but it would constitute an important way to dismiss this account.

Other Broad Concepts of “Empathy”

The latter kind of criticism directed to Preston and de Waal's work can also explain the skepticism for other broad conceptions of empathy. Just to provide a few examples of the kind of definitions I am referring to, let me mention the definitions provided by Simone Shamay-Tsoory, Daniel Batson, and Simon Baron-Cohen.

Empathy is a broad concept that refers to the cognitive as well as the emotional reactions of one individual to the observed experiences of another [...] Emotional empathy may involve several related underlying processes, including, among others, emotional contagion, emotion recognition, and shared pain. On the other hand [...] the term cognitive empathy describes empathy as a cognitive role-taking ability, or the capacity to engage in the cognitive process of adopting another's psychological point of view. This ability may involve making inferences regarding the other's affective and cognitive mental states [...] It seems likely that each component in the empathy network is associated with distinct functions that comprise the empathic response. To fully

characterize the empathy network, it is necessary to identify the roles of each contributing brain region to the processes that support the 2 systems.²⁷

In this passage, the author correctly distinguishes two different systems and provides evidence for their different basic mechanisms and processes. Notwithstanding this distinction, though, she conceives these two distinct systems as pertaining to the domain of the same concept – namely “empathy”.²⁸

In *Altruism in Humans* Daniel Batson provides a very interesting and appealing definition of the relation between empathy and altruism. The definitions provided are extremely well-informed and the conceptual distinction is carried out carefully. Yet, when it comes to defining “empathy”, the definition goes as follows:

I shall use the term *empathic emotion* to refer to other-oriented emotions elicited by and congruent with the perceived welfare of someone else. I shall use *empathic concern* and, as a shorthand, *empathy* to refer to *other-oriented emotion elicited by and congruent with the perceived welfare of someone in need* [...] as defined, empathic concern is not a single, discrete emotion but includes a whole constellation. It includes feelings of sympathy, compassion, softheartedness, tenderness, sorrow, sadness, upset, distress, concern, and grief.²⁹

So, besides the extremely relevant differentiation Batson displays between different phenomena that have been understood as internal to empathy itself, his definition remains too broad, since it includes feelings of sympathy, compassion, softheartedness, tenderness, sorrow, sadness, upset, distress, concern, and grief. The last example refers to Simon Baron-Cohen’s definition of “empathy”:

Empathy occurs when we suspend our single-minded focus of attention and instead

adopt a double-minded focus of attention. [...] When empathy is switched off, we think only about our own interests. When empathy is switched on, we focus on other people’s interests too [...] So we can extend the definition of empathy as follows: *Empathy is our ability to identify what someone else is thinking or feeling and to respond to their thoughts and feelings with an appropriate emotion.* This suggests there are at least two stages in empathy: recognition and response. Both are needed.³⁰

Baron-Cohen’s definition is obviously a very broad one containing both a recognition element – for thoughts as well as feelings – and an appropriate emotional response. However, it is one thing to recognize someone else’s emotional state, and another to recognize her thoughts, and a very different one to respond appropriately to either of them. Moreover, Baron-Cohen derives from this extremely appealing account a straightforward conclusion:

The key idea is that *we all lie somewhere on an empathy spectrum* (from high to low). People said to be evil or cruel are simply at one extreme of the empathy spectrum.³¹

This approach resembles the idea I have analyzed in Preston and de Waal’s article of a connection between empathy and helping behavior. The only difference is that while their focus was on specific behaviors, Baron-Cohen’s focus is on the whole character of a person. So that if someone is said to be a cruel person, this is due to the fact that he has no empathy. These cases are what the author defines as the «zero degrees of empathy»,³² to which many of his works are devoted.

The similarity between this derivation of cruelty from a lack of empathy and the connection shown by Preston and de Waal between empathy and helping behavior exposes Baron-Cohen’s account to the same kind of criticisms I have made of Preston and de Waal’s work.

■ Paul Bloom’s “Against Empathy”

A deeper criticism of Baron-Cohen’s account comes from Paul Bloom:

Strong inclination toward empathy comes with costs. Individuals scoring high in unmitigated communion report asymmetrical relationships, where they support others but don’t get support themselves. They also are more prone to suffer depression and anxiety. Working from a different literature on “pathological altruism,” Barbara Oakley notes in *Cold-Blooded Kindness*, ‘It’s surprising how many diseases and syndromes commonly seen in women seem to be related to women’s generally stronger empathy for and focus on others.’ The problems that arise here have to do with emotional empathy – feeling another’s pain. This leads to what psychologists call empathetic distress. We can contrast this with non-empathetic compassion – a more distanced love and kindness and concern for others. Such compassion is a psychological plus. Putting aside the obvious point that some degree of caring for others is morally right, kindness and altruism are associated with all sorts of positive physical and psychological outcomes, including a boost in both short-term mood and long-term happiness. If you want to get happy, helping others is an excellent way to do so. It is worth expanding on the difference between empathy and compassion, because some of empathy’s biggest fans are confused on this point and think that the only force that can motivate kindness is empathetic arousal. But this is mistaken.³³

I believe Paul Bloom’s influential article *Against Empathy*, which appeared in the *Boston Review*, can be interpreted as criticizing a broad account of empathy in which the latter is seen as the “moral good” and as “the only force that motivates kindness”. The criticism

is somehow directed towards an even broader target, than the one that can be found in Preston and de Waal, Shamay-Tsoory, and Batson’s works: it refers specifically to empathy’s role in social justice and moral behavior. It is easy to see, then, how empathy can be a moral good, and it has many champions. Obama talks frequently about empathy; witness his recent claim, after his first meeting with Pope Francis, that ‘it’s the lack of empathy that makes it very easy for us to plunge into wars. It’s the lack of empathy that allows us to ignore the homeless on the streets.’³⁴

In line with the criticism I have put forward, problems arise when we imply that to act in a good way, or to be helpful, is and can only be due to empathy. There is, though, a difference between Bloom’s criticism and my own. His target is the way in which the term empathy is commonly used in social and political debates, while mine is the way in which empathy is portrayed in the scientific and philosophical domains. Besides this difference in the target, a relation can be traced out between the two domains.

On the one hand, data coming from scientific research underscoring the relation between empathy and helping behavior have been interpreted as a thorough explanation of the link, giving the impression to the nonprofessional audience that the matter is settled and we know everything that is relevant for a helping behavior to occur. References to empathy in the social or political debate are often driven by this intuition: since empathy has been proven to be a good candidate for helping and morally good behavior, then to pursue a more empathic world is to pursue a more just and fair world. Unfortunately, things are not so easily assessed. Paul Bloom’s criticisms of this simplistic viewpoint are well directed. For a society to be just and fair we need much more than empathy itself and it is true that empathy has a wide range of limits and biases. Moreover,

the intuition rests on the assumption that from scientific data we can straightforwardly derive what is morally good or wrong. However, the matter is much more complicated, so that such an assumption is both wrong and dangerous.

On the other hand, the common sense usage of the term has certainly influenced those researchers who have considered it more promising to conceive of “empathy” as a broad concept, so that it would be applicable to all of the common sense uses. However, the difference – that I believe has been neglected – is between the extension of the term “empathy” to its related phenomena – such as identification, emotional contagion – and the derivation of moral behavior, morally good consequences from empathy. The former is true in many scientific and philosophical works,³⁵ the latter is more akin to common sense usage. While the difference remains, even though heavily disregarded in the literature, these misconceptions bear reciprocal influences.

Some elements in favor of conceptual clarification

For the purposes of this paper it will be enough to have shown some difficulties in accounts – that offered by Preston and de Waal has been used as the major example – that deny the necessity of a conceptual clarification within this field. If we conflate automatic neural responses and behaviors under the same broad concept, it will be non-rhetorical to ask why and how things can go wrong. Non-helping behaviors need a more accurate explanation. If the link is direct, how is it possible that sometimes it does not work out?

In general, the relation between empirical data and philosophical insight is more complex than just superimposing a model or a theoretical framework on data, or interpreting the latter as providing evidence for one particular theoretical framework. Moreover, I have mentioned that the influence of com-

mon sense usage on scientific research, as much as that of an oversimplification of scientific data on the social and political domain, can be damaging for both sides. As Daniel Batson says in regard to altruism, although his insight can certainly be applied to empathy itself:

Altruism is not an easy topic; there are many conceptual subtleties, inferential complexities, and empirical challenges that cannot be usefully addressed by oversimplification.³⁶

I would like here to summarize the reasons for the conceptual confusion, that have emerged above, and to provide an idea as to how such confusion can be addressed. Four main aspects have been emphasized.

First, as it has been mentioned in discussing Paul Bloom’s article, the relation between empirical research – as it is portrayed for the general public – and the common sense usage of the term “empathy” has caused several problems for attempts to provide a clear and precise definition of the subject matter.

Second, the linguistic issue cannot be ignored. When dealing with conceptual definitions and speakers’ competence in using a specific term, we cannot disregard the fact that there might be significant differences in different languages. I will not look closely into this issue here, but it would be of extreme interest to pursue *the GRID Project*³⁷ – or something analogous – for “empathy”, “sympathy” and other related phenomena in order to see whether speakers’ competencies with these terms vary across cultures and whether a more detailed distinction entails differences in experiencing the very same situations.

To provide an example of the need to have specific words in one’s own vocabulary in order to be able to understand and apply certain conceptual differences, I will consider Adam Smith’s usage of the term “sympathy”. I believe “sympathy” is used, in Adam Smith’s work, in two very different ways, even though the fact that he possessed only

one word in English complicates the whole issue. On the one hand, passages like the following can be found:

When we see a stroke aimed and just ready to fall upon the leg or arm of another person, we naturally shrink and draw back our own leg or our own arm.³⁸

Yet, on the other hand, Smith most of the times seems to suggest a more robust sense of “sympathy”:

Though our brother is upon the rack, as long as we ourselves are at our ease, our senses will never inform us of what he suffers. They never did, and never can, carry us beyond our own person, and it is by the imagination only that we can form any conception of what are his sensations. Neither can that faculty help us to this any other way, than by representing to us what would be our own, *if we were in his case*. It is the impressions of our own senses only, not those of his, which our imagination copy. By the imagination *we place ourselves in his situation*, we conceive ourselves enduring all the same torments, we enter as it were into his body, and become in some measure the same person with him, and thence form some idea of his sensations, and even feel something which, though weaker in degree, is not altogether unlike them.³⁹

It looks like there are two different phenomena Adam Smith is talking about: on the one hand, there is a sort of instinctive ability to share, to feel together with others; on the other, sympathy seems to arise from imagination and from our ability to imagine what would happen, how we would feel if we were that other person. My claim is that these are two very different phenomena, the former is an ability we are somehow and to some extent endowed with, it is immediate and unreflexive; while the latter is a more complex ability to *understand* others, it requires an ac-

tive engagement of the perceiver in order to access it, and the involvement of imagination.⁴⁰ The former is a form of *sharing*, the latter a form of *comprehension* on the basis of which we judge others’ actions and emotional reactions to be appropriate or not.

Moreover, I claim that the former makes it easier to achieve the latter, but it is not necessary to go from the ability to share to that of understanding. Cases of impairment or reduced ability to share should support this claim, in contrast with Simon Baron-Cohen’s account. I cannot go into this issue more deeply here. But why is it that Adam Smith used the word “sympathy” for both these phenomena? It might be the case that he had a broad and comprehensive concept of it, as those analyzed in the previous paragraphs, yet I think the linguistic issue should not be ignored.

“Empathy” is the more recent term in English, entering the vocabulary only in the last century. The credit for translation into English goes to Edward Bradford Titchener, who coined the word empathy to express the difference between *Einfühlung* (in-feeling) and *Mitgefühlung* (with-feeling) [...] which was already in the English lexicon as sympathy.⁴¹

The fact that the word “empathy” was introduced later than “sympathy” in English can help in the interpretation of Adam Smith apparent confusion. I used Adam Smith here only as an example of what can happen when we do not possess sufficient terms and their appropriate diversification.

Third, another reason for the confusion relates to the fact that these phenomena have been studied from various perspectives. As Coplan and Goldie also noticed⁴² there are several research fields in which empathy has become relevant. Moreover, it has been used for very different purposes. Three main purposes can be identified in the literature.

First, it has been used to «gain a grasp of the content of other people’s minds, and to

predict and explain what they will think, feel, and do».⁴³

To this extent, the understanding of empathy relates to the discussion about how we understand others, in particular our possession of such an ability has been interpreted as providing evidence for a *Simulation Theory* of how we engage with the content of others' minds or as an alternative view both to *Simulation Theory* and *Theory of Mind (ToM)*.

Secondly, empathy has been used to understand our moral behavior as evidence for the revival of sentimentalism and an ethics of care, as opposed to a more detached approach to ethics.⁴⁴

Finally, another research field in which the concept of empathy has been extensively used is the interpretation of our engagement with art and works of art.

The problem with these different uses of the term is that when trying to provide a comprehensive account of empathy, one has to deal with definitions that are rich in theoretical assumptions depending on the specific role that these traditions want to underline. This is possibly one reason why providing a definition of empathy that is an umbrella concept seems a viable option. Though, as I have shown in criticizing such accounts, I believe to do so is not the most useful way to advance our understanding of the ability to empathize.

The concept of empathy has been used to consider a broad range of theoretical questions, from diverse perspectives in numerous disciplines. Different aims, different methodologies, and the difficulty in understanding each other's jargon contribute to controversies in defining the topic. The term empathy has been used for decades – in some cases even centuries – in developmental and social psychology, psychiatry, neuroscience, educational research, social sciences, philosophy, health studies, nursing, ethology, literature, and economics just to mention a number of fields. To accommodate all of these perspectives is definitely a complex matter. One way to do it is to understand to which level each

one can contribute. It is to this extent that Marr's distinction proves itself extremely useful, also in respect to understanding of empathy and it constitutes the last aspect I believe is worth mentioning.

David Marr provides an influential description of how we can account for vision, introducing three levels of explanation that can and have been used also for other domains. Firstly, he noticed that «almost never can a complex system of any kind be understood as a simple extrapolation from the properties of its elementary components».⁴⁵

The same can be said for empathy: understanding the basic processes that enable it – e.g. explaining how human brains implement empathy – is not sufficient to account for the aims we are most interested in. The collection of functional data – though relevant to understand implementation – does not tell us anything about the “macroscopic” issues we aimed to explain. They do not provide evidence for empathy's role in ethics or in our comprehension of other's minds.

If one hopes to achieve a full understanding of a system as complicated as a nervous system [...] then one must be prepared to contemplate different kinds of explanation at different levels of description that are linked, at least in principle, into a cohesive whole, even if linking the levels in complete detail is impractical.⁴⁶

The different levels of explanation Marr has in mind are: computational theory, representation or algorithm, and implementation. In order to introduce them, he uses the example of a cash register. The first and more abstract level of explanation is that of a computational theory, and the aim is that of understanding *what* the device does and *why*. Following Marr's example, the cash register sums. The *what*-question is answered by a theory of addition. The *why*-question is answered by considering the reason for choosing that particular operation – addition – instead of other possible options. It is answered

by a series of constraints.

In order that a process shall actually run, however, one has to realize it in some way and therefore choose a representation for the entities that the process manipulates. The second level of the analysis of a process, therefore, involves choosing two things: (1) a *representation* for the input and for the output of the process and (2) an *algorithm* by which the transformation may actually be accomplished. For addition, of course, the input and output representations can both be the same, because they both consist of numbers [...] If the first of our levels specifies what and why, this second level specifies *how*. For addition, we might choose Arabic numerals for the representations, and for the algorithm we could follow the usual rules about adding the least significant digits first and “carrying” if the sum exceeds 9. Cash registers, whether mechanical or electronic, usually use this type of representation and algorithm.⁴⁷

The last level is represented by the implementation, in the case of the cash register it is the specific machine that physically embodies the algorithm. As Marr underlines, the algorithm that a child uses to add two numbers might well be the same used by the cash register, but obviously the implementation is rather different. Some devices might be more

apt for certain algorithms. Thus, when *designing* a program, it might well be that the choice of either the implementation we want to use or the algorithm that best suits our purposes commits us to certain choices at the other level.

This, obviously, holds only when the aim is that of creating a device and not when the aim is that of *explaining* an existing device that was not produced by us – as in the case of both vision and empathy. The Table 1 (see below) summarizes the questions to which each level has to provide an answer.

Once a distinction between different explanatory levels has been proposed, though, a question concerning the relations between them might arise. Marr explains how he believes the three levels relate to each other as follows:

there is a wide choice available at each level, and the explication of each level involves issues that are rather independent of the other two. Each of the three levels of description will have its place in the eventual understanding of perceptual information processing, and of course they are logically and causally related. But an important point to note is that since the three levels are only rather loosely related, some phenomena may be explained at only one or two of them. This means, for example, that a correct explanation of some psychophysical observation must be for-

Table 1. Reproduction of Marr’s Figures 1-4 (see D.C. MARR, *Vision*, cit., p. 25).

Computational theory	Representation and algorithm	Hardware implementation
What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?	How can this computational theory be implemented? In particular, what is the representation for the input and output, and what is the algorithm for the transformation?	How can the representation and algorithm be realized physically?

mulated at the appropriate level. In attempts to relate psychophysical problems to physiology, too often there is confusion about the level at which problems should be addressed.⁴⁸

The three levels are logically and causally related – as mentioned above, when designing a device some constraints from the implementational level might guide the choice for a certain algorithm to be implemented, or it might be the other way around – yet, they are to some extent independent from one another. One can advocate a computational theory of vision without entering details about how it is represented or implemented.

The kind of confusion Marr individuates in his example is the same one I have tried to show when criticizing Preston and De Waal's article. Saying that there is such a confusion between different levels of explanation does not mean that there can be no relation between them, but recommends a better understanding of the possible contribution of each level to a comprehensive analysis of the phenomenon at issue.

Bearing Marr's distinction in mind and trying to apply it to the debate about empathy, different disciplines can contribute to different explanatory levels. Simplifying for the sake of understanding, one can say that neuroscientific data, for instance, will provide an account of the implementational level, of how our brains are wired in order for empathy to occur; psychology, broadly speaking, will provide evidence for the representational level; and philosophy will provide the computational framework to understand such a phenomenon.

In conclusion, given the growing interest in empathy in various research fields, some departmentalization is still occurring. This makes it hard to provide a comprehensive view of "empathy" and a shared definition. I have analyzed attempts to overcome this problematic division. I have considered attempts that I do not think are promising. In order to account for empathy given the different data coming from different fields, the-

se theoretical approaches have provided an umbrella definition for empathy. I have argued that this is not the most useful way to keep all of the perspectives together. While the aim of these research programs is laudable, since they aim at providing a comprehensive explanation of empathy, the solution they provide is not convincing.

The growth of a debate often not only brings attention to the topic, but also entails some misconceptions and oversimplifications. That is what I believe has happened to empathy. For this reason, I dealt with both the common sense usage of empathy and the different uses in specific literatures. Moreover, I suggested that Marr's distinction of three levels of explanation is applicable to empathy and offers a more useful way to provide a comprehensive understanding of empathy without losing each research field's specificity and without providing an umbrella definition.

Notes

¹ See A. COPLAN, P. GOLDIE, *Empathy. Philosophical and Psychological Perspectives*, Oxford University Press, New York 2011.

² See, e.g., S.D. PRESTON, F.B.M. DE WAAL, *Empathy: Its Ultimate and Proximate Bases*, in: «Behavioral and Brain Sciences», vol. XXV, 2002, pp. 1-20; S.D. PRESTON, F.B.M. DE WAAL, *Empathy: Each is in the Right – Hopefully, not All in the Wrong*, in: «Behavioral and Brain Sciences», vol. XXV, 2002, pp. 49-72; J. RIFKIN, *The Empathic Civilization*, Tarcher/Penguin, New York 2009; S.G. SHAMAY-TSOORY, *The Neural Bases for Empathy*, in: «The Neuroscientist», vol. XVII, n. 1, 2011, pp. 18-24; S. BARON-COHEN, *The Science of Evil: On Empathy and the Origins of Cruelty*, Basic Books, New York 2011.

³ K. STUEBER, *Empathy*, in: E.N. ZALTA (ed.), *The Stanford Encyclopedia of Philosophy*, Spring 2014 Edition, <<http://plato.stanford.edu/archives/spr2014/entries/empathy/>>.

⁴ P. BLOOM, *Against Empathy*, in: «The Boston Review», 2014, <<http://www.bostonreview.net/forum/paul-bloom-against-empathy>>.

⁵ See S.D. PRESTON, F.B.M. DE WAAL, *Empathy: Its Ultimate and Proximate Bases*, cit., p. 2.

⁶ *Ivi*, pp. 5-16.

⁷ See D. MARR, *Vision*, 1982.

⁸ S.D. PRESTON, F.B.M. DE WAAL, *Empathy: Its Ultimate and Proximate Bases*, cit., p. 4, italics in original.

⁹ *Ibidem*.

¹⁰ See S.D. PRESTON, F.B.M. DE WAAL, *Empathy: Its Ultimate and Proximate Bases*, online available, see <http://cogprints.org/1042/1/preston_de_waal.html>.

¹¹ R.A. HINDE, *Emotion: The Relation between Breadth of Definition and Explanatory Power*, in: «Behavioral and Brain Sciences», vol. XXV, n. 1, 2002, pp. 37-38, here p. 38.

¹² S.D. PRESTON, F.B.M. DE WAAL, *Empathy: Its Ultimate and Proximate Bases*, cit., p. 4.

¹³ See M. HOFFMANN, *How Automatic and Representational is Empathy, and Why*, in: «Behavioral and Brain Sciences», vol. XXV, n. 1, 2001, pp. 38-39.

¹⁴ S.D. PRESTON, F.B.M. DE WAAL, *Empathy: Each is in the Right – Hopefully, not All in the Wrong*, cit., p. 54.

¹⁵ S.D. PRESTON, F.B.M. DE WAAL, *Empathy: Its Ultimate and Proximate Bases*, cit., p. 6.

¹⁶ *Ivi*, p. 3.

¹⁷ See F.B.M. DE WAAL, *The “Russian Doll Model” of Empathy and Imitation*, in: S. BRATEN (ed.), *On Being Moved: From Mirror Neurons to Empathy*, John Benjamins, Amsterdam-Philadelphia 2007, pp. 35-48.

¹⁸ *Ivi*, pp. 287-288.

¹⁹ *Ivi*, p. 287.

²⁰ S.D. PRESTON, F.B.M. DE WAAL, *Empathy: Its Ultimate and Proximate Bases*, cit., p. 1.

²¹ See J.H. MASSERMAN, S. WECHKIN, W. WILLIAM TERRIS, “Altruistic” Behavior in Rhesus Monkeys, in: «The American Journal of Psychiatry», vol. CXXI, 1964, pp. 584-585.

²² S.D. PRESTON, F.B.M. DE WAAL, *Empathy: Its Ultimate and Proximate Bases*, cit., p. 1.

²³ *Ivi*, p. 6.

²⁴ S.D. PRESTON, F.B.M. DE WAAL, *Empathy: Its Ultimate and Proximate Bases*, cit., p. 4.

²⁵ See J.D. GREENE, *Why are VMPFC Patients more Utilitarian?: A Dual-process Theory of Moral Judgment Explains*, in: «Trends in Cognitive Sciences», vol. XI, n. 8, 2007, pp. 322-323; J.D. GREENE, R.B. SOMMERVILLE, L.E. NYSTROM, J.M. DARLEY, J.D. COHEN, *An fMRI Investigation of Emotional Engagement in Moral Judgment*, in: «Science», vol. CCXCIII, n. 5537, 2001, pp. 2105-2108.

²⁶ I am grateful to Simone Gozzano and Giovanni Tuzet for their comment on a previous version of this section (presented at the Italian Society for Analytic Philosophy (SIFA), *Philosophy, Analysis and Public Engagement*, University of L’Aquila, September 5th, 2014).

²⁷ S.G. SHAMAY-TSOORY, *The Neural Bases for Empathy*, cit., pp. 18-19.

²⁸ As shown in her Figure 1, see *ivi*, p. 18.

²⁹ C.D. BATSON, *Altruism in Humans*, Oxford University Press, New York 2011, p. 11.

³⁰ S. BARON-COHEN, *The Science of Evil*, cit., pp. 15-16, italics in original.

³¹ *Ivi*, p. 15, italics in original.

³² *Ibidem*.

³³ P. BLOOM, *Against Empathy*, in: «The Boston Review», September 10th, 2014, available at <<http://www.bostonreview.net/forum/paul-bloom-against-empathy>>.

³⁴ *Ibidem*.

³⁵ See S.D. PRESTON, F.B.M. DE WAAL, *Empathy: Its Ultimate and Proximate Bases*, cit.; S.D. PRESTON, F.B.M. DE WAAL, *Empathy: Each is in the Right – Hopefully, not All in the Wrong*, cit.; S.G. SHAMAY-TSOORY, *The Neural Bases for Empathy*, cit.

³⁶ C.D. BATSON, *Altruism in Humans*, cit., p. 6.

³⁷ See *The Swiss Center for Affective Sciences* (<http://www.affective-sciences.org/GRID>).

³⁸ A. SMITH, *The Theory of Moral Sentiments or An Essay Towards an Analysis of the Principles by which Men Naturally Judge Concerning the Conduct and Character, first of their Neighbours, and Afterwards of Themselves*, Henry G. Bohn, London 1759, p. 5.

³⁹ *Ivi*, pp.4-5, italics mine.

⁴⁰ For a similar issue see A. COPLAN, P. GOLDIE, *Empathy. Philosophical and Psychological Perspectives*, cit., p. xi, fn. 10.

⁴¹ J.E. ESCALAS, B.B. STERN, *Sympathy and Empathy: Emotional Responses to Advertising Dramas*, in: «Journal of Consumer Research», vol. XXIX, n. 4, 2003, pp. 566-578, here p. 567.

⁴² A. COPLAN, P. GOLDIE, *Empathy. Philosophical and Psychological Perspectives*, cit., pp. ix-xlvii.

⁴³ *Ivi*, p. ix.

⁴⁴ *Ibidem*.

⁴⁵ D.C. MARR, *Vision*, cit., pp. 19-31.

⁴⁶ *Ivi*, p. 20.

⁴⁷ *Ivi*, p. 23.

⁴⁸ *Ivi*, p. 25.