When do we empathize?

Frédérique de Vignemont

When do we empathize?

Frédérique de Vignemont

Institute of Cognitive Science, 67 bd Pinel, 69675 Bron cedex, France

Abstract. According to a motor theory of empathy, empathy results from the automatic activation of emotion triggered by the observation of someone else's emotion. It has been found that the subjective experience of emotions and the observation of someone else experiencing the same emotion activate overlapping brain areas. These shared representations of emotions (SRE) could be the key for the understanding of empathy. However, if the automatic activation of SRE suffices to induce empathy, we would be in a permanent emotional turmoil. In contrast, it seems intuitively that we do not empathize all the time and that far from being automatic, empathy should be explained by a complex set of cognitive and motivational factors. I will provide here a new account of the automaticity of empathy, starting from a very simple question: when do we empathize? We need to distinguish clearly the activation of SRE and empathy. I will provide a model that accounts both for the automaticity of the activation of SRE and for the selectiveness of empathy. As Prinz says about imitation, the problem is not so much to account for the ubiquitous occurrence of empathy, but rather for its notorious non-occurrence in many situations.

2006 Empathy and Fairness. Wiley, Chichester (Novartis Foundation Symposium 278) p 180–195

According to a traditional view of the mind, we only have an indirect access to what the other thinks or feels through observation and inference. The discovery of mirror neurons in monkeys activated both during action observation and action execution has challenged this view and opened a new pathway for the understanding of intersubjectivity. All we need to do is to exploit one's own resources in order to simulate or recreate someone else's mental states in oneself from a first-person perspective. Functional brain imagery has been recently seeking evidence of overlapping brain activations between feeling and observing the same emotion. Until now, the neural basis of the following emotions and bodily sensations has been shown to be shared: disgust, fear, anger, sadness, happiness, pain, touch (e.g. Calder et al 2000, Carr et al 2003, George et al 1996, Gur et al 2002, Jackson et al 2005, Kesler-West et al 2001, Keyser et al 2004, Phillips et al 1998, Singer 2006, Wicker at al 2003).

Feeling an emotion and observing someone else displaying the same emotion activate the same cortical representation. These shared representations of emotions (SRE) could be the key to the understanding of empathy. We share the same

emotion with someone else because the observation of her emotion triggers automatically the activation of the representation of this emotion from a first-person perspective. However, if the automatic activation of SRE sufficed to induce empathy, we would be in a permanent emotional turmoil. In contrast, it seems intuitively that we do not empathize all the time and that far from being automatic, empathy should be explained by a complex set of factors. I will here provide a new account of the automaticity of empathy, starting from a very simple question: when do we empathize?

The automaticity of empathy

A colleague feels deeply jealous of me because the head of the department decided to send me to a conference that he wanted to attend. Do I share his feeling of jealousy? How could I feel jealous of myself? I do not empathize with him even if I may understand his reaction and feel sorry for him. However, this seems incompatible with recent experimental results about emotions. Brain areas dedicated to subjective experiences of emotions and bodily sensations are activated when observing someone else experiencing the same emotion or sensations whatever the kind of stimulus that is used. It does not seem to matter whether subjects see an isolated body part being injured (Jackson et al 2005) or a facial expression of an unknown person (Adolphs 2002). It does not make a difference whether the study emphasizes the context inducing the sensation (Botvinick et al 2005) or the specific body location injured (Avenanti et al 2005). In all cases, they found shared representations of emotions and sensations that are automatically activated. By automatic, we mean that the activation of SRE is (1) systematic, (2) independent from the context and (3) without the need of any further triggering condition. These results argue in the direction of a bottom-up theory of empathy: a small amount of information of low level is sufficient to induce an empathetic activation, which would be automatic. As Preston & de Waal (2002, p 4) say:

> '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 autonomic and somatic responses, unless inhibited.'

We may better understand their claim, shared by many in the neuroscience of empathy, if we come back to action observation and imitation. From the very beginning, the notion of empathy has been linked to actions. Theodor Lipps suggested that by internally imitating a facial expression, we have direct access to the emotion that triggers this facial expression. The existence of mirror matching systems was considered as a neural evidence of Lipp's theory. Gallese (2001), one of the leaders of this view defends what he calls the 'shared manifold' hypothesis.

He claims that empathy and mirror neurons are just two different levels of description of the same phenomenon of intersubjectivity. Action representations are automatically activated during action observation, even if the movement is not performed by conspecifics (e.g. monkey, human or dogs), as long as it belongs to the motor repertoire of the observer (Buccino et al 2004). The perception of someone else moving suffices to elicit the mental simulation of the performed movement. Unless inhibited, this motor simulation does not remain off line and is physically executed. Imitation is a prepotent response tendency. Indeed, subjects make more errors and are slower to perform a movement when they watch an incongruent movement (e.g. they move their index finger while seeing the little finger moving) (Brass et al 2000). Even if movements observation interferes with action execution, still most of the time we do not imitate other people. Imitation is thus automatic, even if most of the time inhibited. One may then suggest that empathy is not different from imitation. They depend both on shared representations between self and other. They are both automatic. They both remain offline if inhibited. Consequently, several authors have provided what they call 'a motor theory of empathy'.

We have to distinguish between two interpretations of the motor theory of empathy. According to a strong version, you recognize the emotion of others through motor imitation (Gallese 2001, Carr et al 2003, Leslie et al 2004). Empathy is automatic because motor imitation is automatic. According to a weaker version, action should be viewed just as a model of understanding. Both actions and emotions involve representations shared between self and others. They obey the same kind of principles. However, it does not mean that empathy is motoric, even if empathy may share many features with imitation (Preston & de Waal 2002, de Vignemont 2004).

I will not argue here pro or against any of these versions of the motor theory of empathy. I will rather analyse one of the claims that they both make about the automaticity of empathy.

The limits of the motor theory of empathy

How far can we draw the parallel between empathy and imitation? There are at least four main differences. First, it seems that we cannot help but sharing someone else's sadness. In contrast, we imitate because we want to, in order to learn for instance. Second, empathy has a salient phenomenological dimension. I empathize with you if I subjectively experience the same emotion as you. Goldman (1995) describes it as an 'online simulation'. It is difficult to make sense of what empathy would be if it remains offline. In contrast, the study of motor imagery has provided evidence of offline imitation. Third, autism and psychopathy are sometimes described as deficits of empathy, but as far as I know, there is no

pathological case of the reverse, that is, patients that would compulsively empathize all the time with everybody. In contrast, some patients with frontal lesion are no longer able to inhibit and compulsively imitate others (Luria 1966, Lhermitte 1986, Brass et al 2003). Fourth, many factors influence when we feel empathy. For instance, we empathize more with people we feel close to or people we think are fair, as shown by Singer (2006, this volume). In contrast, we may imitate everybody.

Let me pursue further on this latter feature of empathy. At the beginning, we saw that it is difficult to empathize for a subject-directed emotion (e.g. jealousy or angriness toward the empathizer). There are other cases also that raise difficulties for the automaticity of empathy and that show the complexity of the factors influencing when we feel empathy. Imagine that you witness a mother very upset with her son Peter because he made a silly joke to his younger brother Jack who could have been hurt. There are several scenarios of how Peter may react. (a) He regrets what he did and cries. (b) He does not feel sorry, Jack deserved what he got and nothing bad anyway really happened. (c) He does not regret because he did not do anything. His mother is mistaken and he feels her reaction as unfair. With whom do you empathize? Intuitively, in (b) we feel empathy with the mother. In contrast in (c), we empathize with Peter and we feel his mother's behaviour unfair. In (a), the situation is more ambiguous. On the one hand, we may empathize with Peter, the crying little boy. On the other hand, we may empathize with the worried mother.

If we assume that empathy is automatic, then we would have to empathize for two contradictory emotions in all the scenarios, a consequence that goes against our intuitions. One could then reply that the activation of one of the emotions inhibits the activation of the other. Then the question is why this emotion rather than that one. Do we choose with whom we empathize? It does not seem so. The fact that we challenge the automaticity of empathy does not imply that empathy is a voluntary process. We suggest only that empathy is not systematic and needs further additional factors to happen. Preston & de Waal (2002) acknowledge that different factors influence when we empathize, like the familiarity effect. Interestingly, all the factors they describe explain why we feel empathy in some cases, rather than why we do not feel empathy most of the time. Most of the literature about empathy has focused on the conditions that trigger empathy. However, if the activation of SRE sufficed to automatically induce empathy, then there would be no need for any further necessary conditions to explain why and when we feel empathy. As Prinz (2002) says about imitation, the problem is not so much to account for the ubiquitous occurrence of empathy, but rather for its notorious non-occurrence in many situations. The automatic activation of shared representations of emotions cannot be the whole story about empathy.

D2

A two-step model

I would like to suggest here that the problem arises from a confusion in the literature between empathy (sharing the emotional feeling) and SRE (sharing the cortical representation of the emotion). I will now try to provide a model that accounts both for the automaticity of the activation of SRE and for the selectiveness of empathy.

SRE and empathy

Interestingly, a shift in the studies about emotions happened recently. A number of studies have investigated how we recognize the emotions in others based on facial expressions without appealing to our own feelings: subjects have merely to categorize without experiencing the displayed emotion. Indeed, when I watch a face showing fear, I do not feel afraid. There is no empathy involved here. Nonetheless, brain imagery show activation of SRE. In contrast, recent studies about pain have emphasized the subjective phenomenology experienced by subjects while watching the others. When I watch someone being hurt by a needle, I feel almost as if that was happening to me. There is a salient phenomenological dimension of the first-person perspective. In this sense, there is empathy.

As Wicker et al (2003) notice, the strong version of the motor theory of empathy leads to a 'cold hypothesis', which merely requires sharing the facial motor representation of the emotion, in contrast with the 'hot hypothesis' that actually requires sharing the conscious feeling of the emotion. These two views of emotion recognition should not be confused, nor should we reduce empathy to the mere activation of SRE. By definition, empathy involves the subject's emotional experiences and we need to take into account this phenomenological dimension.

We need to draw a sharp distinction between different levels of sharing of emotions. At a primary level, the observation of someone else's emotion triggers the activation of SRE. This activation is automatic and is not inhibited. It occurs independently of the context. It underlies the recognition of the emotion displayed. But the activation of SRE does not necessarily lead to the phenomenological experience of the emotion and can remain offline. At a secondary level, one has consciously access to the emotion associated with the activation of SRE. It is only then that one experiences the emotion of other. It is only then that we can talk of empathy. Far from being automatic, empathy depends on several contextual factors. Consequently, empathy does not need to be inhibited all the time, it is rather sometimes triggered by external conditions. The default rule is not that we empathize with everybody.

I will now turn on the different factors that mediate the transition between the primary and the secondary level, between SRE and empathy. I would like to

suggest that there are at least two main mechanisms that are involved: the distinction between self and other and the evaluation of the emotional event.

Distinction between self and others

By definition, SRE encode both one's own emotions and emotions of others. They do not specify whose emotions they represent, mine or yours. SRE are intersubjective. The activation of SRE expresses this primary lack of differentiation between the self and others. The lack of differentiation implies the necessity of disambiguating the representations by articulating who the subject is (de Vignemont 2004, Decety & Jackson 2004).

In this sense, shared representations of emotions are similar to shared representations of actions and we can go further in the parallel between emotion and action. The activation of mirror neurons does not suffice by itself to determine who is moving, because their content does not specify the agent. That is why we need an additional mechanism that enables us to self-attribute our own actions: the 'Who' system (de Vignemont & Fourneret 2004). Interestingly, this mechanism is also involved in the inhibition of imitation (Brass et al 2005). In contrast with other inhibition mechanisms that are involved for instance in the Stroop task, the inhibition of imitation activates the anterior fronto-median cortex and the temporal-parietal junction, which are both known to be involved in the sense of agency and in perspective-taking. Brass and colleagues claim that the distinction between internally generated and externally triggered motor representations plays a key role to prevent us to imitate someone else's movements. Put it another way, I do not imitate your movements because they are yours and they do not match my own intentions.

Similarly, I would like to suggest that I do not empathize with your emotions because they are yours and they do not match my global feelings and my emotional situation. A crucial requirement for the conscious experience of the emotion would thus be the distinction between my emotions and your emotions. If one detects that SRE are activated following the observation of someone else's emotion, then the activation of SRE does not lead to the phenomenological experience of the emotion in oneself. Indeed, why should one feel what the others feel? One can recognize the emotions of others based on SRE without having to experience them. The offline simulation of emotions suffices, there is no need for empathy. In contrast, if SRE are activated following the experience of an emotional event for oneself, then the activation of SRE leads to the phenomenological experience of the emotion in oneself. The distinction between self and others thus makes the difference between emotional experience in oneself and emotion recognition in others.

If this is true, then we would never empathize. However, we do empathize even when we do not necessarily want to. We need a further step to explain why we feel

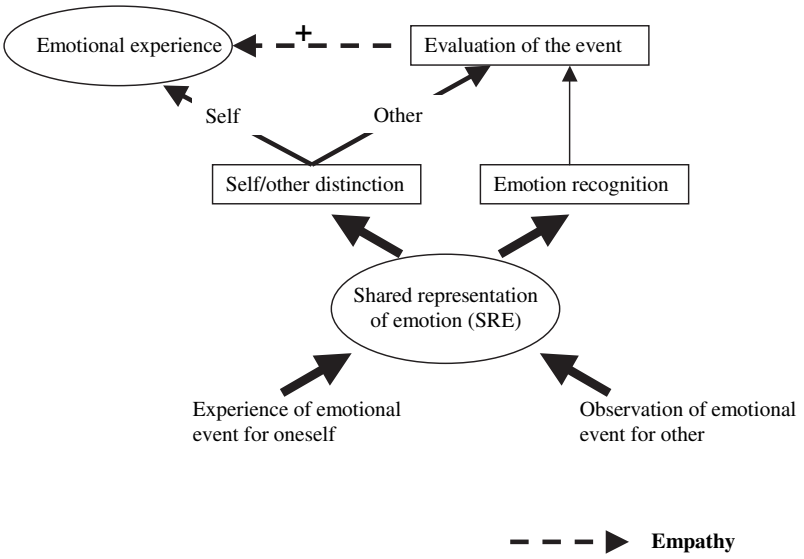

FIG. 1. A two-step model of empathy.

someone else's emotions despite the fact that they are not our own emotions (see Fig. 1).

A set of complex factors

Our hypothesis is that SRE are automatically activated in any emotional context, whether one is at the core of this context or someone else. If the subject of the emotion is the self, then the activation leads to an emotional experience. If the subject of the emotion is someone else, then there is no such experience, unless if other factors reinforce the activation of the SRE leading then to an emotional experience despite the fact that it's someone else's emotion. In this latter case, there is empathy. I would like now to review some of these factors that counterbalance the inhibition of SRE (see Table 1).

The evaluation of the emotional event is not performed voluntarily and remains implicit. The evaluation focuses on three poles: the emotion itself, the person who experiences the emotion and the empathizer.

The first main factor concerns the emotion that one shares with the other. We need again to distinguish between different dimensions. First, some emotions are easier to share than others, like for instance, sadness or pain. I would like to suggest that basic emotions are easier than complex ones and negative easier than positive. I also mentioned at the beginning the impossibility of empathizing for an emotion

D2

TABLE 1 Main factors influencing when we empathize

Type of emotion	*Person feeling the emotion*	*Empathizer*
Target of the emotion	Familiarity	Gender
Complexity	Attitude toward the person	Level of attention
Valence	Similarity and identification	Emotional context
Emotional repertoire		
Saliency and intensity		
Justification		

directed toward the empathizer, like jealousy. Second, the shared emotion has to belong to one's own emotional repertoire (also called the effect of past experience by Preston and de Waal). If you don't suffer from vertigo, you can hardly empathize with me when I am frightened by the void below me. Similarly, the role of motor familiarity for mirror neurons has been demonstrated (Calvo-Merino et al 2005). Third, the shared emotion has to be salient. One does not feel empathy for weak emotions, but rather for strong ones that capture our attention. Four, the shared emotion has to be consistent with the internal and external background. According to a simulationist approach, we put ourselves in someone else's shoes based on the simulation of the mental states of the person and of the context. For instance, can we empathize with someone who starts suddenly screaming and crying with no obvious reason? I predict that we would be surprised rather than share her state of distress. In our previous example, when the mother is unfair with Peter (c), it is difficult to share her anger because we know it is not justified.

The relationship between the empathizer and the subject is also important. This relationship can be understood in three ways. First, there is the familiarity effect as described by Preston & de Waal (2002): we empathize more with relatives or people that we know well. That could be easily explained if we assume a simulationist approach: the more we know about the other, the easier it gets to put oneself in their shoes. Second, there is the emotional attitude that the empathizer has for the subject. That's what Tania Singer shows in her study: men empathize less with people they think unfair. Third, there is the similarity effect, also pointed out by Preston and de Waal. We empathize more with people we can identify to. Then I can really feel the same emotion as if I were you. The dimmer the boundary between the self and other, the easier it is to go beyond this boundary.

A third factor is the overall personal context of the empathizer. According to Baron-Cohen & Wheelwright (2004), women have a higher empathy score that men. More importantly, we are not open to others all the time, paying attention to what they feel. When all our needs are satisfied, we are more likely to empathize (Hoffman 1975). For instance, happy children empathize more (Strayer 1980). D2

Other factors may also play a role and need to be experimentally investigated. Furthermore, each factor may provide contradictory response and will have to be pondered differently. That will decide at the end whether we feel empathy or not.

Conclusion

The discovery of SRE has opened a new pathway for the understanding of empathy but does not suffice in itself for a full account of the complexity of when we empathize. I suggested here that we should distinguish between SRE and empathy. While SRE remains offline, empathy is characterized by the phenomenological experience of someone else's emotion. While the activation of SRE is automatic, empathy is selective. Further work needs to be done to understand the transition from one to the other.

References

Adolphs R 2002 Neural systems for recognizing emotion. Curr Opinion Neurobiol 12: 169–177

Avenanti A, Bueti D, Galati G, Aglioti SM 2005 Transcranial magnetic stimulation highlights the sensorimotor side of empathy for pain. Nat Neurosci 8:955–960

Baron-Cohen S, Wheelwright S 2004 The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism and normal sex differences. J Autism Dev Disord 34:163–175

Botvinick M, Jha AP, Bylsma LM, Fabian SA, Solomon PE, Prkachin KM 2005 Viewing facial expressions of pain engages cortical areas involved in the direct experience of pain. Neuroimage 25:312–319

Brass M, Bekkering H, Wohlschlager A, Prinz W 2000 Compatibility between observed and executed finger movements: comparing symbolic, spatial, and imitative cues. Brain Cogn 44:124–143

Brass M, Derrfuss J, Matthes-von Cramon G, von Cramon DY 2003 Imitative response tendencies in patients with frontal brain lesions. Neuropsychology 17:265–271

Brass M, Derrfuss J, von Cramon DY 2005 The inhibition of imitative and overlearned responses: a functional double dissociation. Neuropsychologia 43:89–98

Buccino G, Lui F, Canessa N et al 2004 Neural circuits involved in the recognition of actions performed by nonconspecifics: an FMRI study. J Cogn Neurosci 16:114–126

Calder AJ, Keane J, Manes F, Antoun N, Young AW 2000 Impaired recognition and experience of disgust following brain injury. Nat Neurosci 3:1077–1078

Calvo-Merino B, Glaser DE, Grezes J, Passingham RE, Haggard P 2005 Action observation and acquired motor skills: an FMRI study with expert dancers. Cereb Cortex 15:1243–1249

Carr L, Iacoboni M, Dubeau MC, Mazziotta JC, Lenzi GL 2003 Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. Proc Natl Acad Sci USA 100:5497–5502

de Vignemont F 2004 The co-consciousness hypothesis: from mirror neurons to empathy. Phenomenology and Cognitive Sciences I3:97–114

de Vignemont F, Fourneret P 2004 The sense of agency: a philosophical and empirical review of the 'Who' system. Conscious Cogn 13:1–19

D2

Decety J, Jackson PL 2004 The functional architecture of human empathy. Behav Cogn Neurosci Rev 3:71–100
Gallese V 2001 The 'shared manifold' hypothesis: from mirror neurons to empathy. Journal of Consciousness Studies 8:33–50
George MS, Ketter TA, Parekh PI, Herscovitch P, Post RM 1996 Gender differences in regional cerebral blood flow during transient self-induced sadness or happiness. Biol Psychiatry 40:859–871
Goldman A 1995 Empathy, mind and morals. In: Davies M & Stone S (eds) Mental simulation: Philosophical and psychological essays, Blackwells, Oxford, p 185–208
Gur RC, Schroeder L, Turner T et al 2002 Brain activation during facial emotion processing. NeuroImage 16:651–662
Hoffman ML 1975 Developmental synthesis of affect and cognition and its implication for altruistic motivation. Dev Psychobiol 11:607–622
Jackson PL, Meltzoff AN, Decety J 2005 How do we perceive the pain of others? A window into the neural processes involved in empathy. NeuroImage 24:771–779
Kesler-West ML, Andersen AH, Smith CD et al 2001 Neural substrates of facial emotion processing using fMRI. Cogn Brain Res 11:213–226
Keysers C, Wicker B, Gazzola V, Anton JL, Fogassi L, Gallese V 2004 A touching sight: SII/PV activation during the observation and experience of touch. Neuron 42:335–346
Leslie KR, Johnson-Frey SH, Grafton ST 2004 Functional imaging of face and hand imitation: towards a motor theory of empathy. NeuroImage 21:601–607
Lhermitte F, Pillon B, Serdaru MD 1986 Human autonomy and the frontal lobes: I. Imitation and utilization behavior. A neuropsychological study of 75 patients. Ann Neurol 19:326–334
Luria AR 1966 Higher cortical functions in man. Basic Books, New York
Phillips ML, Young AW, Scott SK et al 1998 Neural responses to facial and vocal expressions of fear and disgust. Proc Biol Sci 7:1809–1817
Preston SD, de Waal FBM 2002 Empathy: its ultimate and proximate bases. Behav Brain Sci 25:1–71
Prinz W 2002 Experimental approaches to imitation. In: Meltzoff A & Prinz W (eds) The imitative mind: development, evolution and brain bases. Cambridge University Press, Cambridge, MA
Singer T 2006 The neuronal basis of empathy and fairness. In: Empathy and fairness. Wiley, Chichester (Novartis Found Symp 278), p ••–••
Strayer E 1980 A naturalistic study of empathic behaviors and their relation to affective states and perspective skills in preschool children. Child Dev 51:815–822
Wicker B, Keysers C, Plailly J, Royet JP, Gallese V, Rizzolatti G 2003 Both of us disgusted in My insula: the common neural basis of seeing and feeling disgust. Neuron 40:655–664

DISCUSSION

Van Lange: A nice illustration of empathy is when people are attending movies and empathize with the character to the extent that they start to cry. A specific instance of empathy that struck me was during the first Big Brother reality show in The Netherlands, when one of the participants was looking favourite to leave the house in the next vote. A good friend of mine, who is normal fairly balanced, said he would be willing to pay say 60 Euros if this person could stay in the house: he really empathized with this character. These sorts of emotions are not

D2

conscious: there is no analysis of the situation. They just happen. This doesn't involve a lot of cognitive activity.

De Vignemont: I don't claim that we explicitly and consciously analyse the emotional situation. The top–down influences are not available to the subject. You are aware that you empathize, but you are not aware of the reasons why you empathize. Interestingly, in movies you empathize only with one character. We need to understand 'why this one?' and 'why not all of them?'

Blair: You are making direct reference to the more conscious experience of empathy. The problem with that is that we don't have an experimental model of consciousness, so it is not an experimentally tractable question. One of the reasons I never went anywhere near this sort of description is because I knew I'd never be able to have a computational account of it, at least in the short term. It seems to me to be a difficult path to take. You were also shifting from empathy not being automatic, but the neural response or the shared representation being automatic, but we know that this is not correct. The idea used to be that there was an automatic response to, for example, fearful expressions. This has not held up. The degree to which you have that emotional response is determined by the degree to which you attend to the stimulus that generates it. We could flip your argument and say that, yes, we don't empathize all the time, but this is because we are not looking at the face, hands or other triggers. These attentional phenomena can explain this without any complicated alternative processes being invoked.

De Vignemont: I agree that consciousness is a difficult issue to address. Yet empathy involves by definition a conscious emotional reaction similar to the one displayed by the other person. There is a phenomenological aspect that we cannot get rid of. And I think it can be tractable by analysing different situations. For instance, recognition of facial expressions does not elicit a conscious feeling similar to the expression, while seeing someone being hurt does elicit a conscious emotional reaction. By comparing these two situations, we may better understand what is involved in empathy. With regards to your second point, in Tania Singer's experiment, they paid the same amount of attention whether the 'victim' treated the subject fairly or unfairly. Attention cannot explain why she got different results. I don't think we can explain everything by attention, even if it is of course an important factor.

Blair: There are nice models of what attention is about. The Desimone and Duncan model gives a great definition of representational priming leading to attention to particular features of the visual array, driving what the percept is (Desimone & Duncan 1995). Facial expressions are much more powerful than you would anticipate. There is a huge social referencing literature showing that all you need to do is have a novel object in the room, the child is in the room with the mother, looks at the new object, looks at the mother, sees the emotional response of the mother and this determines how the child will respond to the object for

ever more. Susan Mineka has very equivalent monkey data (Mineka & Cook 1993).

De Vignemont: I don't say that we are not using facial expressions, just that they don't elicit a strong conscious experience. That is what we are supposed to have in empathy.

Warneken: Your process model started out with the person's observation of the other's emotion and situation, and then went into the shared representation of emotion. Later on you had an arrow going to interpreting or analysing the situation. How much do the first appraisal of the situation and the later analysis differ? Or should this be construed as some kind of feedback loop?

De Vignemont: When you perceive a sensation, you just have for instance the facial expression of pain. This is the first level. At this level, you do not take into account who is in pain or why. This is just the brute observation of pain. It is only at the later stage of the analysis of the emotional event that you process the whole context surrounding this pain. This processing will be influenced by your folk psychology and your folk moral (e.g. children have to be protected), by other beliefs and desires that you have, by your mood and so on. The first level suffices to elicit the shared representation of emotion, but empathy requires taking the context into account.

Warneken: It is not clear to me that the interpretation of the situation comes only later. You could start out with this. Researchers like Doris Bischof-Köhler use this to distinguish between emotional contagion and empathy proper. When the source of information is the facial expression it is more likely that it is personal distress and emotional contagion, versus when it comes through an inference of the situation where it is more likely to be empathy. The self–other differentiation also has to come into the equation, but the first step is already important.

Gergely: There are some potential complications. You have enumerated a set of conditions which, if they are fulfilled, you feel empathy. This may be so. But what would happen if you have a bad day and every five minutes those conditions are satisfied? I don't think you can feel repeatedly, frequently empathic for a long time. Are there further modulating conditions? It is nice that you have pointed out there is no compulsive empathy as a pathological condition. But I have noticed in my family certain older ladies sit in front of the television crying at frequent intervals.

De Vignemont: Perhaps old ladies would be the equivalent of compulsive imitation for empathy! I agree that we cannot repeatedly empathize with different people, but I think we can keep empathizing with the same person over the course of the movie, for example.

C Frith: I have a vague recollection that there are patients who you can manipulate to laugh or cry uncontrollably just by telling them stories.

Moll: Patients with pontine lesions can manifest pathological crying or laughing.

D2

C Frith: But it isn't clear that this is quite the same as empathy.

Gallese: Part of the analysis you made is very helpful, because it helps in pinning down conditions of activation. It is always a challenge to confront our scientific results with philosophers like you. You are helping us in downplaying our enthusiasm, because as soon as we think we have solved a big problem you tell us that it isn't so big. I learned today that the hardest problem is to explain why we don't empathize all the time. I have some doubt that this heavy reliance on the self-conscious notion of what is going on can be used for pinning down what empathy really is. I am trying to find a minimum level of consensus between your idea of empathy and mine. Would you claim that in order to have empathy, a shared representation of the emotion is a necessary but not sufficient condition?

De Vignemont: Yes.

Gallese: So what is missing from this? Is it the selective activation of this mechanism? This mechanism is by default active all the time. To make the activation of this shared representation of emotion the neural equivalent of what empathy is, then what is missing is the condition of activation. I found some problems when you contrasted the voluntary control of imitation with the apparent automaticity of empathy. You said we can't voluntarily control empathy.

De Vignemont: I agree that there is a kind of paradox here. On the one hand we say that imitation is automatic, but we can control it. On the other hand I say that empathy is not automatic, and we cannot control it! This paradox underlines that imitation and empathy follow different principles. Imitation is inhibited most of the time, but sometimes we can voluntarily release it. It is more difficult to control empathy because there is no inhibition that we can just release. To induce empathy, we need the presence of several factors, and we cannot control all of them. Empathy needs to be triggered while imitation just needs to be released. To go back to your first point, I remember in one of your papers you related mirror neurons with empathy (Gallese 2001). I remember you saying that at the phenomenological level we have empathy but at the neural level we have mirror neurons. I think you agree with me that there is something going on at the phenomenological level.

Gallese: My point was that we should keep different levels of description distinct. We shouldn't imbue neurons with intentional properties. They are just fatty bags letting ions come and go. There is no intentional behaviour in a neuron—even a mirror neuron!

Blakemore: In response to your question about whether there are patients who over-empathize, we found a recent case where this occurs. She's not a patient, though; nor is she an old lady. She is a completely normal healthy friend of ours who feels touch when she sees other people being touched. For example, if she sees someone else being touched on their face she feels it on her face as if she is

D2

being touched. She has always had this and thought it was completely normal. The way we found her was that I was giving a talk about touch and its perception, and whether this kind of person could exist. She raised her hand and asked whether this wasn't completely normal. We studied her and did an imaging study of how her brain is activated by the observation of touch. We found that her mirror system for touch is overactive. She also feels pain that she observes. She has real problems with horror movies.

C Frith: So that's why she is not a nurse!

De Vignemont: That is very interesting. Empathy is a lot about emotions, and for touch the emotional component is very poor. Pain is more interesting because it is at the borderline between emotion and sensation.

Blakemore: There is a distinction between the automatic empathy for pain which doesn't involve you consciously feeling any emotion or sorrow for the person, and empathy where you cognitively put yourself in the person's shoes. She doesn't report doing this.

Molls: Does she feel the same for good and bad characters?

Blakemore: Yes, it is a bottom–up process.

Singer: The attentional thing doesn't account for everything. In the last experiment I did, the modulation of empathy experiment, subjects were equally attending to the fair and unfair person receiving painful stimulation. The experimental condition was exactly the same. The only difference there was their past history with them and their value judgement about these two players.

Blair: You are talking about the manipulation of whether you liked or disliked the person. The straight attention to the stimulus appeared to be identical, but you got a difference between a strong CS association with a much more rich sensory experience for someone you liked rather than someone you didn't. Therefore you have a more boosted signal that activates a stronger emotional response. I wouldn't have explained your data in attentional terms at all.

Singer: If there would have been more much more rich sensory experience for someone you liked this effect should have been controlled by the fact that we are subtracting pain and no pain stimulation for each actor. Thus, your argument does not work here either. Another thing. Why did you say there is no inhibition in empathy? If you could do a time-course analysis with fMRI, you'd want to see whether there is a shared activation of for example pain or touch and then a second re-appraisal process which modulates this activity. It doesn't even have to be top–down inhibition. In my data, I had this dorsolateral prefrontal cortex activity more in men than women when comparing empathic responses to the pain of unfair versus fair players. This activity might reflect modulation of empathic pain responses given men had less of these empathic responses in ACC and AI than women. I don't want to do this claim yet because I would have to design a study specifically to study the Interaction between DLPFC and anterior

D2

insula for example using new methods such as Dynamic causal modelling. But in principle, you could do this kind of experiment to answer the question, and then perhaps you wouldn't have to be as radical as you are. Familiarity, affective link and all these potentially modulatory factors for empathy will have to be explained.

De Vignemont: You are right that there is no temporal dimension in my model. Maybe there is indeed a feedback loop that goes back to the shell representation of emotion and activates it more or less. That could explain some of the results. However, we cannot account for empathy with a purely bottom-up process; we need the top–down input.

Frank: I don't think this was a big part of your case against automatic empathy, but you made a remark about what happens when we see two conflicting emotions in people. Your assessment reminded me of how an economist would look at it: there is a utility function, we have good things and bad things happening, and we just take the net effect, so you are either happy or sad, not both at once. The subjective well-being writers seem to say that this is not the way the happiness and sadness mechanism works. You can experience a happy emotion and a sad emotion at the same time.

C Frith: It seems a pity that consciousness was dismissed. In the imaging work we have no idea whether we are looking at emotional contagion or empathy. We don't know whether it is the conscious or unconscious bit. I don't immediately see how you could separate them out. It would be interesting to study people known as alexithymic, who experience emotions but are not conscious of them in these sense that they don't know what emotion they or having, or even that they are experiencing an emotion at all (Aleman 2005). It would be interesting to know whether these people show empathy. Do they show emotional contagion? The autonomic physiological components of the emotions they experience are larger than normal. There is a suggestion that by being aware of our emotions, things get damped down. This might be part of the mechanism needed for empathizing: you are controlling your own emotions to switch on the one that you think is appropriate to the situation. If real empathy has to be conscious it will be extremely difficult to study it with brain imaging because we should always see the emotional contagion. At the beginning you said that nurses can't be experiencing everything because it would be terrible for them, but by the end it seemed you were saying that they would get all the emotional contagion.

De Vignemont: There would not necessarily be emotional contagion in nurses, but rather an activation of the shared representations of emotion system. Even this activation may be less strong, as noticed by Avenanti et al (2005), who had a nurse among their subjects. She showed a reduced empathetic activity. Perhaps she was habituated to the display of pain.

D2

References

Aleman A 2005 Feelings you can't imagine: towards a cognitive neuroscience of alexithymia. Trends Cogn Sci 9:553–555

Avenanti A, Bueti D, Galati G, Aglioti SM 2005 Transcranial magnetic stimulation highlights the sensorimotor side of empathy for pain. Nat Neurosci 8:955–960

Desimone R, Duncan J 1995 Neural mechanisms of selective visual attention. Annu Rev Neurosci 18:193–222

Gallese V 2001 The 'shared manifold' hypothesis: from mirror neurons to empathy. J Consciousness Stud 8:33–50

Mineka S, Cook M 1993 Mechanisms involved in the observational conditioning of fear. J Exp Psychol Gen 122:23–38