



*University of Milano Bicocca,*

*Faculty of Psychology*

*Ph.D. in Experimental psychology, linguistics and cognitive neuroscience*

*The neuropsychology of empathy for  
pain: how social differences affect the  
empathic brain.*

Supervisor: Professor Marcello Gallucci

Candidate: Matteo Forgiarini

*Ai miei genitori e ad Angelo.*

*Al professor Marcello Gallucci, grazie.*  
*To professor Mahzarin Banaji, thank you.*

# Index

Index .....	4
Introduction.....	8
Empathy: theoretical background .....	9
Stereotypes, prejudice and discrimination .....	15
Implicit measures: a way to assess implicit attitudes .....	19
The implicit measures in the implicit social cognition theoretical background	22
Neural correlates of empathy .....	27
The pain matrix .....	27
The empathy for pain.....	30
Empathy and social neuroscience .....	44
Racism and the empathy for pain on our skin .....	56
Introduction.....	56
Experiment 1: Racism and indifference to the pain of other people. ....	58
Method .....	58
Measures .....	59
Subjects.....	60
Results.....	61
Experiment 2: implicit attitudes and empathy for pain.....	62
Measures .....	62
Subjects.....	63
Results.....	63
Experiments 1 and 2: general discussion.....	66
Experiment 3: race and pain perception, an experimental study .....	71
Introduction.....	71
Method .....	71
Measures .....	73
Subjects.....	73
Results.....	73
Discussion.....	80
Racism and empathy for pain: an fMRI experiment.....	83
Introduction.....	83
Stimuli.....	84

Method.....	85
Behavioural measures.....	86
fMRI data acquisition.....	86
fMRI analyses.....	86
Subjects.....	88
Results.....	89
Behavioral results.....	89
fMRI results.....	89
Discussion.....	98
Religion and empathy for pain: a pilot study.....	101
General discussion.....	106
References.....	111

Socrates: *Imagine a number of men living in an underground chamber, with an entrance open to the light, [...] in which they have been confined, from their childhood, [...] they are obliged to sit still and look straight forwards, because their chains render it impossible for them to turn their heads round: and imagine a bright fire burning some way off, above and behind them, and an elevated roadway passing between the fire and the prisoners, [...]*



Glaucon: *I have it, he replied.*

Socrates: *Also figure to yourself a number of persons walking behind this wall, [...].*

Socrates: *They resemble us, [...]. For let me ask you, in the first place, whether persons so confined could have seen anything of themselves or of each other, beyond the shadows thrown by the fire upon the part of the chamber facing them? Certainly not, if you suppose them to have been compelled all their lifetime to keep their heads unmoved. And is not their knowledge of the things carried past them equally limited?*

Glaucon: *Unquestionably it is.*

Socrates: *And if they were able to converse with one another, do you not think that they would be in the habit of giving names to the objects they saw before them?*

Glaucon: *Doubtless they would.*

Socrates: *Again: if their prison-house returned an echo from the part facing them, whenever one of the passers-by opened his lips, to what, let me ask you, could they refer the voice, if not to the shadow which was passing?*

Glaucon: *Unquestionably they would refer it to that.*

Socrates: *Then surely such persons would hold the shadows of those manufactured articles to be the only realities.*

Glaucon: *Without a doubt they would.*

Socrates: *Now consider what would happen if the course of nature brought them a release from their fetters, [...] and that the dazzling splendor renders him incapable of discerning those objects of which he formerly used to see the shadows. What answer should you expect him to make, if someone were to tell him that in those days he was watching foolish phantoms, but that now he is somewhat nearer to reality, and is turned toward things more real, and sees more correctly; above all, if he were*

*Categories engender meaning upon the world, like paths in a forest". [...] "an ethnic prejudice is a category concerning a group of people, not based on defining attributes primarily, but including various noisy, possibly false, attributes, leading to disparagement of the group as a whole". (Allport, 1954).*

*to point out to him the several objects that are passing by, and question him, and compel him to answer what they are? Should you not expect him to be puzzled, and to regard his old visions as truer than the objects now forced upon his notice?*

Glaucou: *Yes, much truer.*

*(Plato, 514-520 b.c.)*

English translation from

[http://www.age-of-the-sage.org/greek/philosopher/myth\\_allegory\\_cave\\_plato.html](http://www.age-of-the-sage.org/greek/philosopher/myth_allegory_cave_plato.html)

*Mopping-up is what normal science is all about! This paradigm-based research is "an attempt to force nature into the pre-formed and relatively inflexible box that the paradigm supplies". No effort is made to call forth new sorts of phenomena, no effort to discover anomalies. When anomalies pop up, they are usually discarded or ignored. Anomalies are usually not even noticed and no effort is made to invent a new theory (and there's no tolerance for those who try).*

*(Kuhn, 1962)*

Since when I started studying social psychology I have thought about these texts of Plato and Kuhn and I have wondered how much we are different from men in the Plato's cave or from scientists that neglect data not consistent with already known theories. Social psychologists have been always studied the difference between shadows and reality. During my Ph.d. I tried to investigate whether and how races create shadows in our (social) reality.

# Introduction

The present doctoral thesis aims at exploring the relationship between the social context and the brain functioning. On the base of the social neuroscience tradition, my work wants to investigate whether and how social cues affect how our brain and mind work and process emotional stimuli.

Specifically this project focuses on the effect of race and religion on the empathy for other people pain. Researchers who aim to play a role in social neuroscience filed wish to understand the bindings between what we learned from the world and how our brain and body work. Do stereotypes, prejudices and first person experiences affect how we react to the pain of other people? Does other people's pain always trigger our body arousal and cognitive reaction? What kind of social cues are significant for my mind and my brain to decide whether to react to other people's emotions? Are we always aware of these cues as well as the role they play in our everyday life?

These questions guided my work, my research project and the experiments I conducted.



## **Empathy: theoretical background**

Imagination enables us to project ourselves into the place of other persons, experiencing sensations that are generally similar to, although typically weaker than, those of the other person (Adam Smith, philosopher and economist, 1759/1976).

Although the etymology is clear – we know that it comes from the ancient Greek *empathia* (passion), which is composed of “en” (in) and “pathos” (feeling) and then it turned into the German *Einfühlung* (“feeling into” something) – the word “empathy” has nowadays a very complex meaning. Perhaps a lay person could state that empathy is the ability of understanding a perceived or imagined feeling state of another human being. This is not likely to be sufficient to describe what empathy is. Even scientists do not completely agree on what empathy is.

According to Singer and colleagues (Singer & Lamm, 2009), empathy occurs when an observer perceives or imagines someone else’s (i.e., the target’s) affect and this triggers a response such that the observer partially feels what the target is feeling. In other words, empathy is the capacity to share and understand the emotional states of others (Batson, Duncan, Ackerman, & Buckley, 1981). According to De Vignemont and Singer (De Vignemont & Singer, 2006) there is empathy if: (i) one is in an affective state; (ii) this state is isomorphic to another person’s affective state; (iii) this state is elicited by the observation or imagination of another person’s affective state; (iv) one knows that the other person is the source of one’s own affective state. In the present work, I will refer to that definition when I use the word empathy.

The definition is complete and relevant because it allows us to shed light on the differences between relevant key concepts and components generally associated with the broad concept of empathy, such as mimicry, emotional contagion, sympathy, and compassion. Although these concepts are very alike one another and they usually occur simultaneously, mimicry is defined as the tendency to automatically synchronize affective expressions, vocalizations and movements with those of another person (Hatfield, Cacioppo, & Rapson, 1994). Furthermore, this definition of empathy allows distinguishing empathy from other related phenomena. Cognitive perspective-taking, for example, does not meet the first condition. One can always be able to understand mental states of others, including affective states, but can do that without being in an affective state. Similar to empathy, sympathy refers to an affective state related to the other. However, it does not meet the condition of isomorphism (e.g. I feel sorry for you if you are in pain, but I'm not in pain as well), so that it shouldn't be taken as a synonymous. Finally, emotional contagion involves affect sharing but does not meet the condition of self-other distinction (e.g. the baby starts crying because other babies cry but the baby is not necessarily aware that the other is the source of their affective state).

Interestingly, the condition of self other distinction in order to be in presence of empathy, is one of the most important in empathy literature and it has been widely debated. In social and developmental psychology literature, classic researches claimed that empathy involves identification of self with the other showing that perceived similarity increases empathic emotion (Krebs, 1975) and describing empathy as “a feeling of oneness” and “mutual identification”. Hornstein in 1978 (Hornstein, 1978) listed three conditions under which such

identification can arise: When the other's welfare promotes one's own welfare, when self and other are linked by similarity, and when self and other share common membership in a social category or group.

In a similar vein other researchers claimed that the other for whom empathy is felt becomes included in the self. Wegner in 1980 (Wegner, 1980) suggested that “empathy is one way in which positive forms of social behavior may be motivated by selfish desires” (p. 131). He believed this was so because empathy involves an “extension of self” to include the other (p. 132). Empathic feelings “stem in part from a basic confusion between ourselves and others” (p. 133); when we feel empathy we “consider others as though they were ourselves”.

Describing the consequences of this inclusion, (Aron & Aron, 1986) (p. 29) proposed the following: As P [Person] includes more and more aspects of O [Other] into P's self, P comes in a sense to include O - not just aspects of O - into P's self. That is, P feels as much or nearly as much satisfaction when O is satisfied, or pain when O is hurt, as P would if these had happened to P. P plans for O's happiness and welfare as if it were P's. P “identifies” with O, or is even in some sense “united” with O.

Cialdini and colleagues (Cialdini, Brown, Lewis, Luce, & Neuberg, 1997) reported research designed to test a self-expansion explanation of the empathy-helping relationship. They found that when effects of self-reported oneness (measured by perceived self-other overlap and rated appropriateness of the term "we" when speaking of the other) were partialized, the association between self-reported empathy and willingness to help disappeared. These results led the authors to propose that empathic concern and oneness are both influenced by a crucial feature of relationship closeness: perspective taking. When one feels empathic concern, it is normally due to the perspective taking that attends

relationship closeness and that leads to self-other overlap. Upon experiencing empathic concern for another, then, an individual is consequently informed of a likely degree of oneness with that other, and prosocial action is more probable as a result (p. 491).

The importance of perceived oneness and the role it plays in the links between perspective taking, empathic concern (i.e. altruistic motivation) and helping, is well supported and described by results reported by (Maner et al., 2002). The authors tested the link between empathic concern and helping both by using an experimental perspective-taking paradigm already used to demonstrate empathy-associated helping and by assessing the empathy-helping relationship while controlling for different non altruistic motivations. The data were consistent with a model in which experimentally increasing oneness should increase helping, whereas instructing participants to take the target's perspective should only increase helping when participants have no explicit information about oneness. At the same time the source of altruistic motivation - empathic concern – showed a significant correlation with helping. Indeed, the zero-order relationship between empathic concern and helping was nearly as strong as the relationship between perceived oneness and helping. But most important, the empathy-helping relationship was not statistically significant when nonaltruistic motivators were partialized. Finally the data indicated that helping was functionally mediated by only nonaltruistic constructs (perceived oneness, nonempathic negative affect) and not by empathic concern.

On the other hand, there are psychologists who claim that empathy is not “oneness driven”. By claiming that self-other distinctiveness is a necessary condition for the empathy-helping relationship they support the empathy-altruism hypothesis (Batson, Early, & Salvarani, 1997; Batson, 1991). This hypothesis

defines empathy as an other-oriented emotional response congruent with the other's perceived welfare, it defines altruism as a motivational state oriented to increase the other's welfare, and it contrasts altruism with egoism, a motivational state with the ultimate goal of increasing one's own welfare. According to the empathy-altruism hypothesis, the conditions that evoke empathy increase concern for the other's welfare but do not reduce self-other distinctiveness. Empathy involves awareness of the other's particularity and uniqueness.

Coherently with this hypothesis, Batson and colleagues' findings (D. Batson et al., 1997) showed that individuals induced to feel high empathy perceived much the same distinction between themselves and the target of empathy as did individuals induced to feel low empathy. They used three different measures of merging: perceived similarity, Aron's IOS scale, and overlap of perception of self and other on personal attributes. Although there is clear evidence that manipulation of perceived similarity can affect empathy (Krebs, 1975) they found no reliable evidence of the reverse. Data did not support the hypothesis according to which participants in the high empathy condition helped more because they became psychologically indistinguishable from the other and experienced what they perceived the other was experiencing, or that they confused self and other or considered the other as self, or that they expanded the self to include the other or lessened the self-other distinction (Aron & Aron, 1986).

Taking together these data seem to indicate that empathy for another person does not imply that the source of empathy has to believe the target is alike him/her, or even the target is him/her. On the contrary, this hypothesis suggests that the source has to think that out there that is someone with he/she wants to connect with and to share feelings and emotions.

During my research I investigated mainly how similarity between the perceiver and the target of empathy affects the empathic response. Race plays a main role in determining similarity between two persons.

# Stereotypes, prejudice and discrimination

Impressions that are similar... especially if a label is attached... tend to cohere into categories (generalization, concepts). All categories engender meaning upon the world. Like paths in a forest. They give order to our life-space... The principle of least effort inclines us to hold to coarse and early formed generalizations as long as they can possibly be made to serve our purposes. (Allport, 1954) (pp. 175,176)

The first author to investigate stereotypes and prejudice from a modern point of view is likely to be Gordon Allport.

Beside him, researchers studied stereotypes and prejudice on an individual level. The authoritarian personality theory (Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950), the subtle racism model (Pettigrew & Meertens, 1995) and the dissociation model (Devine, 1989) hinged on people's conflicts between internal states (desires and beliefs) and the social.

Allport (Allport, 1954) developed the nascent ideas the social categorization is driven by the context. In *The nature of prejudice*, Allport claimed the stereotyping, prejudice and hence discrimination come from cognitions in social context.

According to the author, humans inevitably categorized objects and people in their world, then to prejudice is fully normal. Like with tables, apples or houses, people categorize other persons into ingroups and outgroups, loving one and, Allport thinks, hating the other. Through the comparison of the target attributes and a relevant mental representation, perceivers attempt to establish some degree of fit

between the target and the representation, with a bias toward determining subcategory membership. Once satisfactory fit is established, this representation will guide the processing of information about and behavior toward the target. If fit is sufficiently poor, perceivers will individuate the target, essentially developing a unique representation. In a similar vein, Fiske and Taylor (Fiske & Taylor, 1991) expressed a similar belief "Once a person is categorized as Black or White, male or female, young or old, the stereotypic content of the schema is likely to apply regardless of how much or how little the person looks like the typical category member" (p. 121).

In physical and social world, "categories engender meaning upon the world, like paths in a forest". In this view, "an ethnic prejudice is a category concerning a group of people, not based on defining attributes primarily, but including various noisy, possibly false, attributes, leading to disparagement of the group as a whole" (Allport, 1954).

In 1981, based on Allport's theory, Tajfel proposed (Tajfel, 1981) that prejudice results from the need for a positive social identity with an ingroup that can think the outgroup as a devaluated contrast.

According to this theory, mere perception of belonging to different groups triggers ingroup favoritism (minimal group paradigm) (Tajfel, 1986; Tajfel, Flament, Billig, & Bundy, 1971) and outgroup discrimination. This means that if people are assigned to groups on the base of variable with no social, economic or personal valence, they will favor members of the same group merely because they are members of the same group. In other words, Tajfel's researches showed that we tend to prefer and to favor people we think they are like us just because we are alike.



Hence, why do people categorize? To shed light (and meaning) in the world, as Allport said, and to minimize within group differences and to accentuate between group differences. Hence, categorization process is always and deeply context dependent.

Specifically racial group members whose appearance most closely resembles our representation of the "typical" category member are more likely to be viewed through the lens of the category stereotypes and evaluations. This phenomenon could be termed racial phenotypicality bias.

As Russel pointed out (Russell, Wilson, & Hall, 1992), in the USA White Eurocentric facial features in Blacks were seen as evidence of White ancestry, leading to inferences of racial superiority. After the abolition of slavery, lighter skin provided better social, educational, and economic opportunities.

Then, to the degree that race, gender, sexual orientation, nationality, or political opinion differentiate two clusters of people (they are salient in a certain context), that particular self-other distinction will be used, with all the consistent stereotypes, prejudices and discrimination.

In 1981 Hamilton (Hamilton, 1981) disserted on the cognitive misery of people claiming that people are overwhelmed by complexity of the social environment and forced to conserve scarce mental resources.

Categorization is a good heuristic to overcome cognitive limit as well and to manage and control a huge amount of information with few concepts.

Hence categorization increases the perceived homogeneity of groups members, especially of outgroups members (Wilder, 1981) who are seen as more similar each other than ingroup members see each other.

Notably, after categorizing, people can misperceive outgroup targets according to their implicit personality theories for that group (Ashmore, 1981).

The main point of cognition-in-contest approach to categorization is likely to be the claim that categorizing is a normal and useful cognitive mechanism used by people in almost every condition in real life. From categorization arise both prejudice and discrimination which can be thought of as respectively the affective and behaviorally side of categorization.

Even though we know races do not exist, I will describing my research using the term “race” and interpreting the results on the base of that concept. This choice comes from the kind of stimuli I used in my experiments. As I will describe later, subjects have been presented with very short videos containing Black, Chinese and Caucasian people. Actors, do not speak or show any non verbal communications. I think that on the base of these information subjects can not understand or make inferences on the ethnic groups actors belong to. Subjects can just categorize actors on the base of their genotypic traits, i.e. skin color and facial cues.

In my research I also used a newer type of measures that are supposed to reach internalized associations in subjects’ minds. This type of measures, the implicit measures, are very likely to be more efficient in the social neuroscience field than the explicit measures.

# Implicit measures: a way to assess implicit attitudes

Since psychologists aim at assessing core psychological processes, they are faced with two key problems of *direct* measures (e.g., self-reports), i.e. introspective limits (Nisbett & Wilson, 1977) and susceptibility to self-presentation or socially desirable responding (Paulhus, 1984). The idea that there might be in our mind more than we can tell, we want to tell or we are socially allowed to tell is what triggered the development of indirect measures. Such measures were developed aiming at obtaining measures of internalized associations without having to ask the person directly. After an initial period in which new measures did not demonstrate sufficient reliability and therefore their application to the assessment of implicit constructs at the individual level was highly problematic (Greenwald & Banaji, 1995), three years later the IAT was introduced as the first implicit response-time measure that proved to be reliable. Thus, the great demand for reliable indirect measures helped to make the IAT widely accepted. The IAT's easy applicability could have favored the widespread use in a large number of different topics, starting from psychology.

The Implicit Association Test is thought to assess the strength of associations between target categories (e.g., Black persons vs. White persons) and attribute categories (e.g., negative vs. positive), both supposed to be perceived on bipolar dimensions, by comparing the response latencies for two differently combined categorization tasks. Participants are presented with stimuli that represent the four categories (e.g., names typical for Blacks vs. Whites and negative vs. positive words) and are told to categorize them with the help of two response keys, each assigned to two of the four categories.

The IAT's most important assumption is that if two concepts are highly associated in subjects' minds, *categorization will be easier (then faster)* when the two associated categories share the same response than when they require different responses.

Tab. 1 presents a typical task sequence of the IAT (a racial attitude IAT) with seven blocks, 1st and 2nd and 5th are practice blocks, 3rd and 4th, 6th and 7th are so-called critical blocks.

Block	N of trials	Task	Response key assignment	
			Left key	Right key
1	20	Target discrimination	Black	White
2	20	Attribute discrimination	negative	positive
3	20	Initial combined task	Black, negative	White, positive
4	40	Initial combined task	Black, negative	White, positive
5	20 or 40	Reversed target discrimination	White	Black
6	20	Reversed combined task	White, negative	Black, positive
7	40	Reversed combined task	White, negative	Black, positive

Tab. 1 Typical structure of a race IAT.

During the IAT procedure participants are trained to press a left key for the left category (“Black” in this example) and a right key for the right category (“White” here) in the first block of 20 trials (target discrimination). In the second block of 20 trials, they are trained to press the same left key for “negative” stimuli and the same right key for “positive” stimuli (attribute discrimination). The critical third block (20 trials) and fourth block (40 trials) mix up the attribute and target discrimination. At this point, participants are instructed to respond left to “negative” OR “Black” stimuli, and right to “positive” OR “White” stimuli. In the fifth block (20 or 40 trials), target categorization is inverted and what was on the

right turns on the left. Participants are trained to press the left key for “White” stimuli and the right key for “Black” stimuli. The diagnostically relevant sixth block (20 trials) and seventh block (40 trials) again combine the attribute and the now reversed target discrimination. Participants now have to respond left to “negative” and “White” stimuli, and right to “positive” and “Black” stimuli.

The main point is the difference in performance between the initial combined blocks (third and fourth block) and the reversed combined blocks (sixth and seventh block). *This time difference is the so-called IAT effect.*

The IAT effect is interpreted in its size and direction and it indicates the relative association strength between the target and attribute categories. Talking of a race IAT, individuals with implicit prejudices against Blacks are supposed to respond faster and more accurately when Black stimuli and negative attributes are paired in the same side of the screen and are associated to the same key (and White stimuli and positive attributes to the other key) compared to the reversed configuration (Black and positive are assigned to one key, White and negative to the other key). It is important to note that IAT effects are always a relative effect meaning that it is always a measure of an attitude over a measure of a second attitude. This means that a racial attitude IAT effect does not permit any conclusions about subject’s evaluation of Blacks, but provides only information about an individual’s preference for Blacks over Whites (or Whites over Blacks) (Teige-Mocigemba, Klauer, & Sherman, 2010).

The current standards in IATs designing comprehend (Teige-Mocigemba, Klauer, & Sherman, in press): (a) the instruction to respond as quickly and accurately as possible, (b) correction of erroneous responses as indicated by a “X” below the stimulus, (c) intertrial interval of 150 ms to 750 ms (250 ms may be most often used), (d) five to six stimuli per category (at least two), (e) alternation

between target and attribute stimuli in the combined blocks, (f) otherwise randomized trial order if group differences are the main focus of the experiment, but (g) a fixed random trial order for all participants in correlational studies in order to reduce confounds of procedural and interindividual variance

Since 1998, many scoring procedures have been proposed for calculating IAT effects (Greenwald, Nosek, & Banaji, 2003). The key point in all scoring methods is that the implicit effect descends from the difference between the initial combined blocks (3<sup>rd</sup> and 4<sup>th</sup> block) and the reversed combined blocks (6<sup>th</sup> and 7<sup>th</sup>).

Specifically, most researchers report analyses based on the so-called *D* measures which are optimized with regard to the IAT's psychometric criteria (e.g., internal consistency, high correlations with explicit measures, resistance to some extraneous procedural influences). *D* measures are calculated by using both correct and incorrect responses, with incorrect response latencies being increased by an error penalty, and an individual standardization similar to that in Cohen's effect size measure *d*.

### **The implicit measures in the implicit social cognition theoretical background**

The field of implicit social cognition has been based on integrative theories for a while (Fazio, 1990; Greenwald et al., 2002). These theories stand in the tradition of dual-process models (Chaiken & Trope, 1999) and theorize that two qualitatively different psychological processes account for most of the phenomena observed in implicit social cognition studies. Recently, a number of theories have invoked the notion of psychological systems in order to group and systemize the various processes (Lieberman, Gaunt, Gilbert, & Trope, 2002) (Smith & DeCoster, 2000; Strack & Deutsch, 2004). In essence, these models suggest that

subgroups of processes “correlate” and thus form mental systems (Deutsch & Strack, 2006).

One of the most relevant theories is the Reflective-Impulsive Model (RIM; Strack & Deutsch, 2004). It aims at integrating cognitive and behavioral processes that underlie overt social behavior. The core tenet of the RIM is that social cognition and behavior depend on function of a Reflective System (RS) and an Impulsive System (IS), each operating according to different representations and computations. In this model, the two systems operate interactively and serve different functions.

The IS groups up the associative memory that contains evaluative and semantic associations with the notion of a procedural memory that contains habits. It is also responsible for generating affect. The RS is specialized in generating propositional representations based on what is activated in the IS and, if necessary, to fulfill executive functions such as overcoming habitual responses or putting together action-plans in new situations or when habits fail (Lieberman, 2003). It contributes to create thoughts, judgments and which can lead to verbal- or non-verbal behavior. The RS is the system that generates behavioral decisions and translates them into overt behavior via intending and goal-regulation. In other words the IS rules processes that are not always known by people and hence plays a role in determining behaviors while people are not fully aware of them. RS resembles what have been always thought of as the explicit social cognition, i.e. decisions, thoughts and behaviors people are aware of.

Principles of dual-system models have been used to interpret data in many topics within social cognition research and beyond, including evaluative processes (De Houwer, In press; Gawronski & Bodenhausen, 2006), addiction (Deutsch & Strack, 2005), and consumer behavior (Strack, Werth, & Deutsch, 2006).

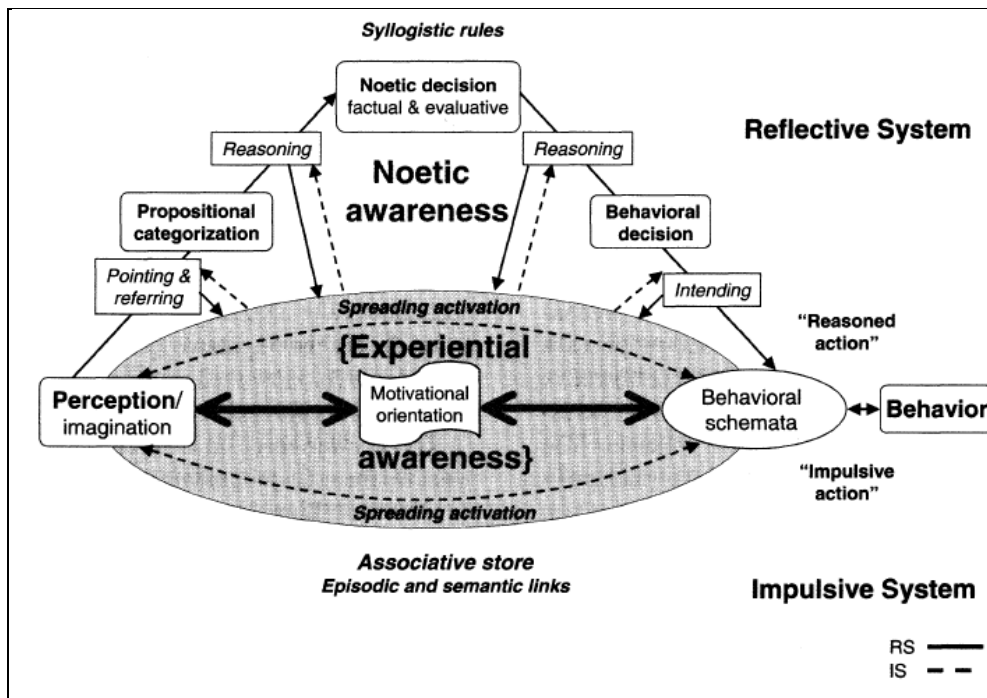


Fig. 1 The figure show a graphical representation of what Reflective System and Impulsive Sytem are and how they work (Strack & Deutsch, 2004).

Even though there are several issues on which the reflective-impulsive model takes a divergent position (Greenwald & Banaji, 1995; Smith & DeCoster, 2000) there is no doubt that the reflective-impulsive model bears a family resemblance to explicit-implicit models in that both notions hold that different mechanisms may mediate between valence and behavior (Strack & Deutsch, 2004).

Moreover Strack and colleagues go one step further and propose to locate implicit processes in the impulsive system, whereas explicit processes are thought to take place in the reflective system. Specifically, authors use the terms explicit versus implicit for psychological processes but not for mental contents. Hence an attitude is defined as a belief following from an evaluative decision that follows from reflection about what is good or bad. On the other hand, evaluative associations are links between concepts and lead to evaluative responses in the



impulsive system. This means that explicit measures capture people's knowledge or beliefs, implicit measures are measures of their associative structures.

Beyond redefining the implicit versus explicit dichotomy, the reflective-impulsive model may help to understand how behaviors may be influenced by explicit and implicit mechanisms. The reflective-impulsive model can account for such influences through its inherent interconnections between conceptual and behavioral representations as well as through its motivational orientations. Many studies (Fazio, 2001) have demonstrated that attitudes may be activated very quickly, efficiently, unintentionally, or even unconsciously on the perception of the attitude object.

The main process through which attitudes may influence behavior spontaneously is by influencing the appraisal of social and environmental cues. If motivation and cognitive resources are high, however, more deliberate considerations and effortful search in memory may prevail and determine the perception of the situation and finally the behavior. Fazio suggested that the activation of such positive and negative associations may tune the perception of the situation, which in turn influences behavior in a spontaneous fashion. The impulsive system goes further and allows for an even more direct path to behavior. Thus, from the perspective of the reflective-impulsive model, the behavioral component of an attitude has a reflective component because it's derived from an evaluation and an impulsive component because it refers to action tendencies that are directly associated with the evaluative features of the attitude object.

Another advantage of the reflective-impulsive model is that it lends itself to approaches from neuroscience (Lieberman et al., 2002). The prediction that reflectively generated attitudes and impulsive responses could rely on different

systems are supported by a neuropsychological study conducted by Phelps and her colleagues (Phelps et al., 2000). Authors showed that amygdala activity was correlated with impulsive behavior indicating prejudice against African Americans, i.e. it was correlated with a race IAT subjects' scores, while it was not correlated with reflective expressions of racial attitudes, i.e. explicit racism scales. This suggests that the amygdala may be capable of detecting stimulus valence extremely fast (Morris, Ohman, & Dolan, 1998) and before it can be processed by the reflective system.

At the same time a recent TMS study (Avenanti, Sirigu, & Aglioti, 2010) using a Blacks/Whites race IAT, showed that despite the lack of explicit racial bias in the sample, participants with higher implicit ingroup preference presented greater corticospinal reactivity to ingroup models over outgroup models' pain. This effect was similarly present in both white and black subjects (see later in the text for a more complete review).

Taking together experimental evidences and theoretical models seem to indicate implicit measures as the most promising kind of measures to explore relations between neurological activities, people's thoughts and internalized associations.

That's way I chose to use implicit measures in my experiments.

My research project is grounded in the social neuroscience field. Next chapters show how explicit and implicit measures and experimental paradigms have been applied to study the neural correlates of empathy.

## Neural correlates of empathy

“A great deal of social psychological research has been devoted to the question which perceptual, affective, and cognitive mechanisms enable us to “put ourselves into someone else’s shoes.” Surprisingly, it took quite some time for the field of neuroscience and in particular of functional neuroimaging to dare to make contributions to this challenging pursuit (Carr, Iacoboni, Dubeau, Mazziotta, & Lenzi, 2003; Morrison, Lloyd, Di Pellegrino, & Roberts, 2004; Singer et al., 2004; Wicker et al., 2003). This might be attributed to the complexities inherent in this multidimensional psychological phenomenon as well as to the methodological challenges of bringing such an idiosyncratic and context-dependent phenomenon into a scientific environment that requires well-controlled and reproducible experiments.” (Singer & Lamm, 2009)

### The pain matrix

The so-called pain matrix is a network of regions that process information in circuits that can broadly be assumed to process the affective, sensory, cognitive, motor, inhibitory, and autonomic responses stimulated by a noxious event (Derbyshire, 2000).

One of the first studies of brain function during painful experience revealed activation in the thalamus and anterior cingulate, primary and secondary somatosensory cortices (Talbot et al., 1991). Since then, there have been many studies that shaped what is now called “the neuromatrix for pain” (Davis, Taylor, Crawley, Wood, & Mikulis, 1997; Derbyshire et al., 1997; Svensson, Minoshima, Beydoun, Morrow, & Casey, 1997). This body of research showed that pain matrix expands well beyond the thalamus and anterior cingulate, and primary (S1)

and secondary (S2) somatosensory cortices to include the midbrain region of the periaqueductal gray (PAG) and the lenticular complex as well as the insula, orbitofrontal (Brodmann's area [BA] 11,47), prefrontal (BA 9,10,44–46), motor (BA 6, Supplementary motor area, and M1), inferior parietal (BA 39,40), and anterior cingulate (BA 24,25) cortices (ACCs) (See Fig. 2).

Study	Method	n	Gender	Stimulation	Reported regions											
					PAG	Tha	LN	Cau	Ins	OFCx	PFCx	MCx	39/40	S1	S2	
<b>Tonic heat pain</b>																
Becerra et al. [2]	fMRI	12	M	46° C	—	C†	—	—	C†	—	C†	C†	B†	C†	B†	
	fMRI	12	M	46° C	—	C†	—	—	C†	—	C†	B†	B†	C†	B†	
Tolle et al. [3]	PET	12	M	1° C > threshold	†	C†	I†	—	C†	I†	I†	—	—	—	—	
Derbyshire and Jones [4]	PET	12	M		—	I†	C†	C†	I†	—	I†	—	C†	—	C†	
Svensson et al. [5]	PET	10	?	45° C to 49° C	—	C†	I†	—	B†	—	—	C†	C†	C†	C†	
Adler et al. [6]	PET	9	4 F	47° C to 48° C	—	I†	?	—	?	I†	L†	C†	—	—	—	
Rainville et al. [7]	PET	8	3 F	↑ Unpleasantness 47° C	†	B†	—	—	C†	—	—	C†	—	C†	C†	
Craig et al. [8]	PET	11	3 F	47° C	—	C†	C†	—	C†	—	—	—	—	C†	C†	
<b>Tonic cold pain</b>																
Kwan et al. [9]	fMRI	13	6 F		NA	—	NA	NA	NA	NA	NA	B†	NA	NA	NA	
Petrovic et al. [10]	PET	10	M		†	—	—	—	B†	—	—	—	—	C†	C†	
Davis et al. [11]*	fMRI	12	3 F		NA	†	—	NA	†	NA	NA	NA	NA	NA	†	
Casey et al. [12]	PET	9	3 F	6° C	—	B†	C†	—	B†	—	—	C†‡	—	C†	—	
<b>Visceral tonic pain</b>																
Ladabaum et al. [13]	PET	15	9 F	Gastric distension	†	B†	B†	B†	B†	?	B†	—	—	—	—	
Mertz et al. [14]†	fMRI	16	14 F	Rectal distension	NA	†	NA	NA	†	NA	†	NA	NA	NA	NA	
Naliboff et al. [15]	PET	12	2 F	Rectal distension	†	B†	B†	R†	L†	B†	B†	B†	L†	L†	—	
Binkofski et al. [16]	fMRI	5	M	Esophageal distension	—	—	—	—	B†‡	—	—	R†	—	B†	B†	
Aziz et al. [17]	PET	8	1 F	Esophageal distension	—	—	—	—	B†	R†‡	B†	B†	—	B†	—	
Silverman et al. [18]	PET	6	?	Rectal distension	—	L†	—	—	R†	—	L†	—	L†	—	—	
<b>Miscellaneous tonic pain</b>																
Iadorola et al. [19]	PET	13	5 F	Capsaicin	†	C†	B†	?	B†	—	—	I†	—	C†	C†	
May et al. [20]	PET	7	M	Allodynia Capsaicin	—	—	I†	?	B†	B†	B†	B†	B†	C†	I†	
Porro et al. [21]	fMRI	24		Subcutaneous ascorbic acid injection	NA	NA	NA	NA	NA	—	†	†	NA	†	NA	
Andersson et al. [22]	PET	6	M	Capsaicin	—	—	—	—	B†	—	I†	C†	—	C†	C†	

Fig. 2 Results of Studies Using PET or fMRI and Experimental Tonic Noxious Stimulation in Control Subjects. \*Regions with  $\geq 50\%$  of subjects showing activation are included in this table, laterality is ignored for simplicity. †Mertz et al. (Mertz et al., 2000) did not assess laterality. ‡Region is included based on inspection by the current author and was not reported by the study authors. B—bilateral; C—contralateral; Cau—caudate; F—female; fMRI—functional magnetic resonance imaging; I—ipsilateral; Ins—insular cortex; L—left; LN—lentiform nucleus; M—male; MCx—motor cortex; NA—not assessed; OFCx; orbitofrontal cortex; PAG—periaqueductal gray;

PET—positron emission tomography; PFCx—prefrontal cortex; R—right; S1—primary somatosensory cortex; S2—secondary somatosensory cortex; Tha—thalamus; 39/40—inferior parietal cortex (BA 39/40). Up arrow indicates increased regional cerebral blood flow or blood oxygenation level dependent signal; question mark indicates uncertainty regarding activation (adapted from) (Derbyshire, 2000).

	Reported regions: studies with activation, %										
	PAG	Tha	LN	Cau	Ins	OFCx	PFCx	MCx	39/40	S1	S2
Tonic pain: controls	33	76	47	22	100	26	65	70	39	68	63
Heat	25	100	62	12	100	25	62	62	50	62	75
Cold	50	50	33	0	100	0	0	66	0	100	66
Visceral	40	67	40	40	100	50	83	60	40	60	20
Other	33	67	33	33	100	25	75	75	33	75	67
Phasic pain: controls	45	81	85	25	90	15	70	67	45	70	70
Phasic pain: patients	33	33	100	0	100	33	66	33	66	33	66
Clinical tonic pain: patients	45	64	27	8	85	67	64	67	50	33	33
Neurogenic	60	67	0	0	67	80	83	60	20	20	0
Visceral	25	75	25	0	100	50	25	50	75	67	75
All controls	40	81	72	24	95	24	64	62	38	71	64
Male	30	75	55	20	90	25	80	65	40	65	70
Female	75	80	100	50	100	25	100	50	50	75	25
All patients	37	53	37	6	78	53	58	53	47	29	35
All males (24 studies)	29	67	54	17	92	29	83	67	42	54	62
All females (7 studies)	80	71	80	20	100	40	86	40	40	60	40

Fig. 3 The Percentage of Studies with Increased Activation in the Regions Selected for Fig. 2. \* The male figures include all the studies with exclusively male subjects or more than 80% male subjects. The female figures include all studies with exclusively female subjects or more than 50% female subjects. The patient groups were intermixed male and female making a gender division impossible. Cau—caudate; Ins—insular cortex; LN—lentiform nucleus; MCx—motor cortex; OFCx—orbitofrontal cortex; PAG—periaqueductal gray; PFCx—prefrontal cortex; S1—primary somatosensory cortex; S2—secondary somatosensory cortex; Tha—thalamus; 39/40—inferior parietal cortex (BA 39/40) (adapted from) (Derbyshire, 2000).

Among other regions, insula primary somatosensory area (S1) and anterior cingulate cortex play a main role in pain perception coding different cues of the pain stimuli. Specifically, insula region has an important role in autonomic regulation: electrical stimulation of the insula in rats, cats, dogs, monkeys, and humans elicits changes in blood pressure, heart rate and respiration. The insula has connections with a large number of subcortical sites involved in autonomic control. It receives afferents directly from the parabrachial nucleus and from the lateral hypothalamic area and a particularly large input from the central nucleus of the amygdala.

The S1 region is well suited to localization and provides detailed information about the location and characteristics of particular noxious stimuli. At the same time it seems is not involved in coding and triggering the general

cognitive or affective responses. Although it is still not fully understood why S1 responses vary during noxious stimulation, subtle differences in the intensity of the pain experience, the type of stimulation, and the relative intensity of nonpain tactile stimulation are likely to be important (Craig, Reiman, Evans, & Bushnell, 1996). Bushnell et al. (Bushnell et al., 1999) have demonstrated substantially greater S1 response in subjects trained to attend to the intensity of the painful stimuli and less activation when directing attention away from the painful stimulus. These data suggest that cognitive and affective processes can influence even localization, a seemingly computationally independent and precise function.

Unpleasant affect is integral to the experience of pain. ACC seems to be mainly involved in cortical representation of the aversive and affective aspects of pain (Singer et al., 2004). Observations of patients with chronic pain following a neurosurgical procedure to cut the white matter tract underlying the ACC (“cingulotomy”) support this view. Although such patients remain aware of their pain and are still able to discriminate noxious stimuli, their pain no longer bothers them (Santo, Arias, Barolat, Schwartzman, & Grossman, 1990).

Then the so-called pain matrix seems to be a network that allows us to perceive a painful stimulus as an amalgam of cognition, affect, and sensation mediated through widespread brain regions.

Social neuroscience aims at shading light on whether this network is affected by social cues of stimuli. That’s what I did with my research project.

### **The empathy for pain**

According to Hein and colleagues (Hein & Singer, 2008) many studies suggest that there are at least two different routes to put us in the shoes (the mind) of the other person (De Vignemont & Singer, 2006; Decety & Lamm, 2006;

Preston & de Waal, 2002). One way is to share the other person's feelings in an embodied manner, commonly referred to as empathy. The other route is to cognitively infer and speculate about the state of the other person, known as "theory of mind", "mentalizing", "mindreading" or "cognitive perspective taking".

The majority of neuroscience studies on neural correlates of empathy have addressed this issue by studying empathy for other people pain (Bufalari, Aprile, Avenanti, Di Russo, & Aglioti, 2007; Cheng et al., 2007; Gu & Han, 2007; Jackson, Brunet, Meltzoff, & Decety, 2006; Lamm, Batson, & Decety, 2007; Lamm, Nusbaum, Meltzoff, & Decety, 2007; Morrison, Peelen, & Downing, 2007). Many studies investigating empathy for pain have shown that empathy for pain is supported by neuroanatomical circuits underlying both affective (empathic) and cognitive processes (Decety & Jackson, 2004; Hein & Singer, 2008; Lamm, Batson et al., 2007; Olsson & Ochsner, 2008).

On one hand, a neural network including bilateral anterior insula (AI) and anterior cingulate cortex (ACC) (Decety and Jackson, 2004; Hein and Singer, 2008; Olsson and Ochsner, 2008) is thought to underlie the affective components of empathy. AI and ACC code the autonomic and affective dimension of pain and in particular, the subjective experience of empathy when perceiving pain or distress in others (Decety and Jackson, 2004; Hein and Singer, 2008; Olsson and Ochsner, 2008). (Singer et al., 2004). On the other hand, studies on perspective taking typically ask participants to take the perspective of a person shown on a cartoon or described in a story. Brain regions activated by cognitive perspective taking include medial prefrontal regions, the superior temporal sulcus (STS), extending into the parietal lobe (temporo-parietal junction), sometimes also the temporal pole (Fig. 4) (Hein & Singer, 2008). Furthermore, cognitive components

of empathy seem to rely on subregions of medial prefrontal cortex (MPFC) (Lamm et al., 2007a), which responds more when people are identifying with others who are more similar to themselves (Mitchell et al., 2006).

As discussed above, *cognitive perspective taking* differs from empathy because cognitive perspective taking does not require sharing of feelings. To be aware of someone else's feeling just needs cognitive understanding of what another person is feeling at that moment. Then, beside theoretical discussion and experimental data, from a neuroscientific perspective, it is important to demarcate empathy from cognitive perspective taking on the basis of different neural networks for empathy and cognitive perspective taking outlined in Fig. 4 (Hein & Singer, 2008).



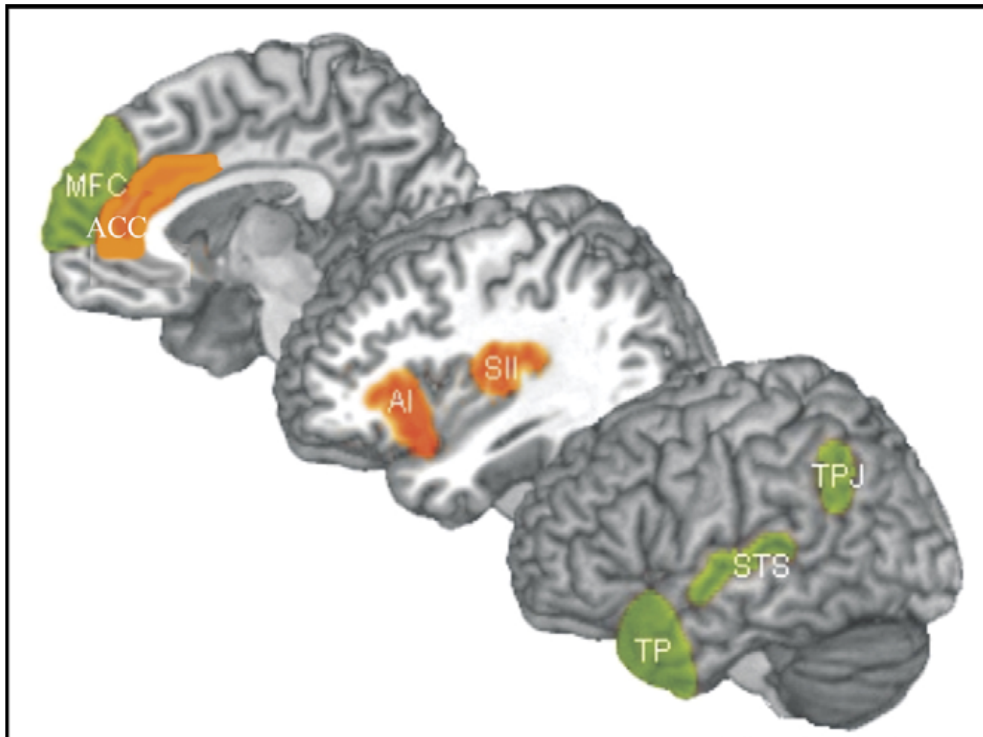


Fig. 4 Schematic overview of brain regions typically involved in understanding others on the basis of cognitive perspective taking (green) and empathy (orange); the latter measured in the domain of empathic brain responses to pain, disgust, taste, and touch. MPC, medial prefrontal cortex; ACC, anterior cingulate cortex; AI, anterior insula; SII, secondary somatosensory cortex; TP, temporal poles; STS, superior temporal sulcus; TPF, temporo-parietal junction (adapted from) (Hein & Singer, 2008).

Since 2002 Preston and de Waal (2002) proposed a neuroscientific model of empathy, one which suggests that observing or imagining another person in a particular emotional state automatically activates a representation of that state in the observer, along with its associated autonomic and somatic responses. Other authors have also suggested that shared neural representations play a general role in understanding other people's mental states.

They claim that shared representations provide us with a simulation of their corresponding sensorimotor, affective, or mental states (Gallese 2003a; Goldman 2006).

Such accounts hold that the capacity to project ourselves imaginatively into another person's perspective by simulating their mental activity using our own

mental apparatus lies at the root of our mature mind-reading abilities, and the reasoning of these accounts has been extended to the domains of actions and feelings: To understand what another person is doing we simulate his movements using our own motor program; to understand what another person is feeling, we simulate his feelings using our own affective programs (see also Keysers&Gazzola 2006). Indeed, this so-called shared representations account of social interaction and intersubjectivity has become the dominant explanation of the hemodynamic activation patterns observed in recent fMRI studies of empathy.

In 2004, Singer and colleagues (Singer et al., 2004) gathered for the first time data on empathy for pain coherent with shared network hypothesis. For the first time the authors measured empathy for pain with an “in vivo” paradigm. According with this paradigm, the subject scanned either can feels pain in first person or he/she is aware of when another subject is feeling pain while the target subject is sitting right beside the scanner.

The scanned subject is presented with a series of visual cues that tell him/her what the other subject is feeling at a specific moment.

By using this paradigm they assessed brain activity while volunteers (women) experienced a painful stimulus and compared it to that elicited when they know that their partner was experiencing pain.

Comparison of brain activity associated with painful and nonpainful trials in the “self” condition revealed increased activity in contralateral SI/MI, in bilateral SII with a peak activation in contralateral posterior insula extending into SII, in bilateral mid and anterior insula, in ACC in right ventrolateral and mediodorsal thalamus, brainstem, and mid and right lateral cerebellum (Fig. 5, A to D, green). Some of areas involved in this network were also activated when pain was applied to the partner, that is, when pain was vicariously felt (Fig. 5, A to D, red).

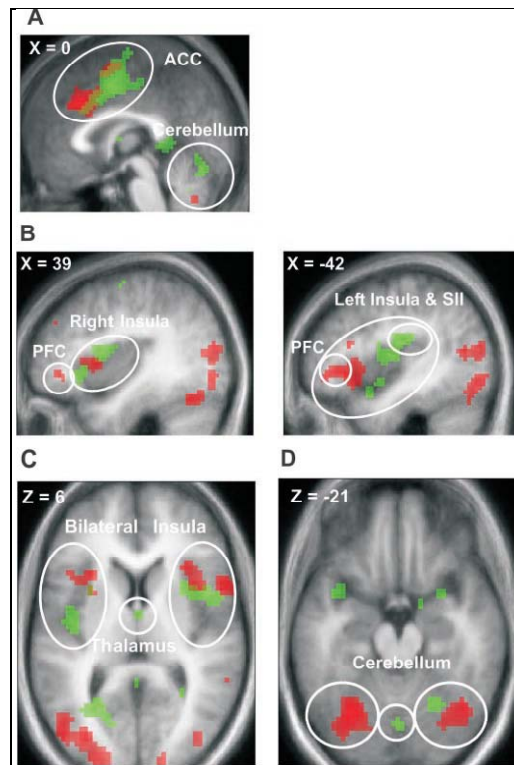


Fig. 5 Pain-related activation associated with either experiencing pain in oneself or observing one's partner feeling pain. Areas in green represent significant activation ( $P < 0.001$ ) for the contrast pain–no pain in the “self” condition and areas in red for the contrast pain–no pain in the “other” condition. The results are superimposed on a mean structural scan of the 16 subjects. Activations are shown on sagittal (A and B) and axial (C and D) slices. (A) Activation in ACC and cerebellum. (B) Bilateral insula cortex extending into lateral prefrontal cortex, left posterior insula extending into secondary somatosensory cortex (SII), bilateral occipital cortex, and fusiform cortex. (C) Bilateral insula and mediadorsal thalamus. (D) Middle and lateral cerebellum/ fusiform gyrus. For coordinates of peak activations from “self” and “other” conditions, see tables S1 and S2 (adapted from) (Singer et al., 2004).

A conjunction analysis and a masking procedure revealed that just ACC (caudal and posterior rostral zones), bilateral middle insula and AI (with a peak activation in the right AI), brainstem, and lateral cerebellum were activated both in fist hand pain and during vicarious pain.

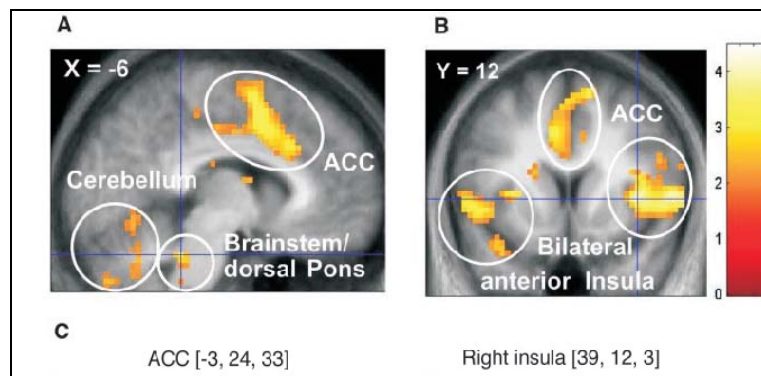


Fig. 6 Shared networks observed when pain was applied to self or to the partner. (A) and (B) illustrate results of a conjunction analysis between the contrasts pain–no pain in the context of self and other at  $P < 0.001$ . Results are shown on sagittal (A) and coronal (B) sections of the mean structural scan. Coordinates refer to peak activations and are in mm (adapted from) (Singer et al., 2004).

Singer and colleagues also showed that individual differences in empathy account for activity in ACC (posterior rostral zone) and left AI (but not right AI). As measured by two empathy scales, the Balanced Emotional Empathy Scale (Mehrabian & Epstein, 1972) and the Empathic Concern Scale (a subscale of the Interpersonal Reactivity Index) (Davis, 1983), ACC and AI significantly covaried with subjects' scores in empathy scales.

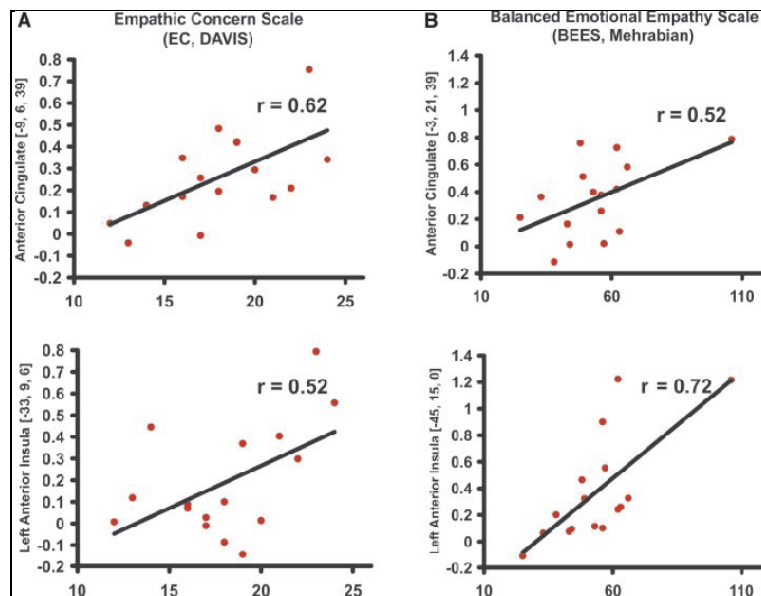


Fig. 7 Activation level (parameter estimates) observed within peaks of the ACC and the left insula during empathy-related conditions (pain–no pain in other) are significantly correlated with individual differences in empathy as measured by (A) the Empathic Concern Scale of Davis and (B) the Balanced Emotional Empathy Scale of Mehrabian. The lines represent the linear best fit;  $r$  refers to the correlation coefficient. All correlations are significant on the  $P < 0.05$  level. Peak activations lie within regions of ACC and left insula that were activated in the simple contrast pain–no pain in others. Coordinates refer to peak activations and are in mm (adapted from) (Singer et al., 2004).

Taking together, these data suggest that empathizing with the pain of others does not involve the activation of the whole pain matrix, but is based on activation of areas involved in affective evaluations of other people pain. In 2004, these data lead the author to argue that, from a functional and evolutionary perspective, people need intensity and location information of a noxious stimuli when it's hurting their own body. Vice versa, the understanding of someone else's emotional reaction to pain just requires a representation of the subjective relevance of the stimulus as reflected in the subjective unpleasantness that the other person feels. These information seem to be all we need to share and understand why and how much another person is in pain.

This initial finding has been replicated and extended using a variety of paradigms and methods. There is now a growing body of literature suggesting that somatosensory processing can also be activated when we witness another person's pain, in particular, when our attention is directed to the somatosensory aspects of the pain experience (Bufalari et al., 2007; Cheng, Yang, Lin, Lee, & Decety, 2008; Lamm, Nusbaum et al., 2007).

Consistently with these findings, Lamm and colleagues (Lamm, Porges, Cacioppo, & Decety, 2008) showed that (contralateral) right primary somatosensory cortex was activated when participants witnessed another person's left hand being pierced. This activation overlapped with individually determined somatosensory representations of touch of the hand that had been determined in a separate localizer run. In a similar vein, a TMS study by Avenanti and colleagues (Avenanti, Paluello, Bufalari, & Aglioti, 2006) showed that motor-evoked potentials are modulated when participants observe a needle deeply penetrate the target's hand, but not when they observe a pinprick. This suggests that somatosensory quality of pain plays a main role in determining whether somatosensory areas will or will not be involved in empathy for pain.

It has to be noted that most of more recent studies on empathy for pain presents subject with pictures or videos of painful unknown faces (Lamm, Batson et al., 2007; Saarela et al., 2007) or body parts in painful situations (Cheng et al., 2007; Gu & Han, 2007; Jackson et al., 2006; Jackson, Meltzoff, & Decety, 2005; Lamm, Nusbaum et al., 2007; Morrison & Downing, 2007; Morrison et al., 2007; Ogino et al., 2007) (See Fig. 8).

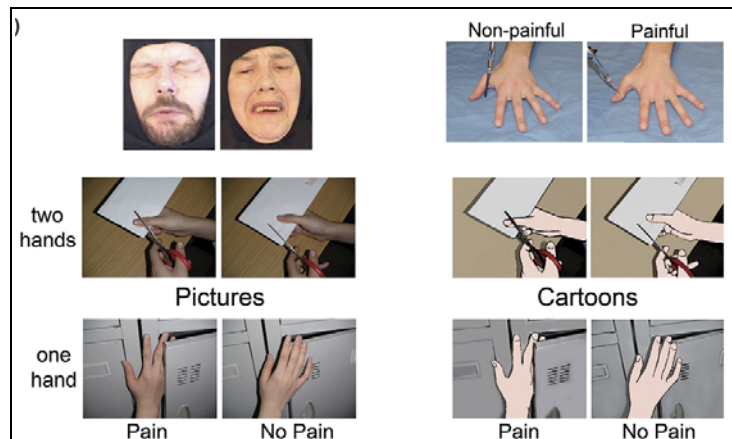


Fig. 8 Results and example stimulus material of empathy for pain studies. Examples of painful faces (Saarela et al., 2007), and pictures and cartoons of body parts in painful situations (top (Lamm, Nusbaum et al., 2007), bottom (Gu & Han, 2007) used to investigate empathy for pain (adapted from) (Hein & Singer, 2008).

This body of research seems to suggest that neural simulation of the pain of another person occurs independently of the affective link between the empathizer and the person in pain. People seem to be able to share or at least to understand other people feelings even if they are unknown, main street men and women.

Does it mean we always automatically empathize with others when exposed to their emotions, regardless of who they are. Is this true? Or are we more selective? As De Vignemont and colleagues suggest (De Vignemont & Singer, 2006), if we were to consciously feel what they feel all the time, we would be in permanent emotional turmoil, leaving no room for our own emotions. The authors suggest that empathy is not merely the consequence of the passive observation of emotional cues but that it is subject to contextual appraisal and modulation.

We fully agree with this thesis on the base of several years of behavioral research in social psychology and social neuroscience fields investigating the role of modulatory factors in empathy, such as similarity, contest-based information and familiarity.

From my point of view, De Vignemont and Singer (De Vignemont & Singer, 2006) proposed maybe the most interesting model that summarizes how and why empathy can be modulated.

I want to use the model proposed by the authors to shed light on the types of factors that can affect empathy and how they are supposed to work.

The authors propose that: a) empathy is modulated by appraisal processes and (b) this modulation occurs even at the subpersonal level of a neural empathic response, and can be fast and unconscious.

Four main types of modulatory factors shape the model (Tab. 2):

Intrinsic features of the shared emotion.

The intensity, saliency and valence (positive versus negative) of the emotion communicated by the target person might affect the intensity of the empathizer's empathic response. Moreover, it might be easier to empathize with primary emotions such as fear, happiness or sadness than with secondary emotions such as jealousy.

Relationship between the empathizer and the target.

In 2006 Singer and colleagues (Singer et al., 2006) found that fairness of target modules empathic responses toward a stranger who interacted with the study subjects. In 1975 Krebs (Krebs, 1975) showed that similarity plays a main role in determining the strength of empathic reactions. Subjects who have been led to think to be similar to targets empathized most with the performer (the target) both when he communicated pleasure and pain. Furthermore familiarity (Cialdini et al., 1997) between the two protagonists and how much protection or care, e.g. nurturance (Batson, 2005) the target needs and whether the emotion is directed towards the empathizer or not (the person in pain being angry or jealous about the empathizer) could also be essential in determining empathic reactions.



### Characteristics of the empathizer

Authors underline that also gender (Baron-Cohen & Wheelwright, 2004; Eisenberg & Carlo, 1995), personality, age (Eisenberg & Morris, 2001) and past experiences of the empathizer could be important. An empathizer who does not suffer from vertigo can barely empathize with a target who is frightened by the void below him because he does not have the specific feeling of vertigo in his repertoire. In such a case, the empathizer might engage in cognitive perspective-taking rather than empathizing.

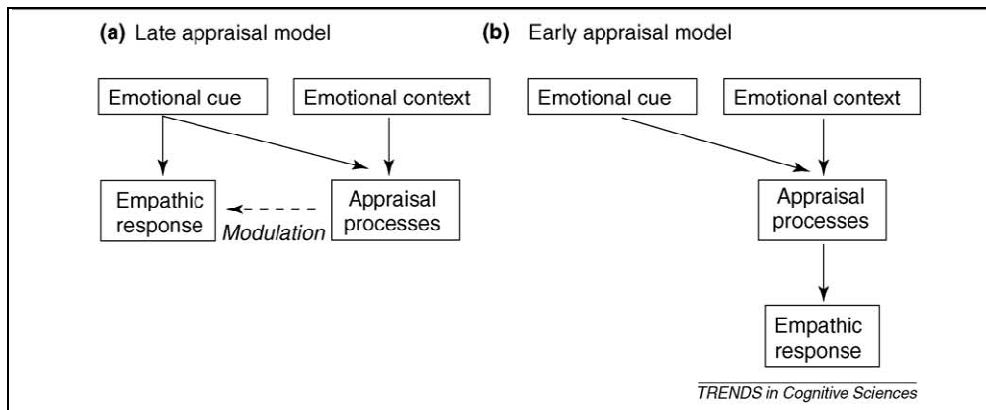
### Situative context

This last category refers to specific situations in which empathizer is relating with strange scenes such as the target starts crying even if he/she is not justified. Or it might be argued that empathizing can also become difficult if the empathizer is simultaneously confronted with two or more targets and the targets are expressing different emotions.

Features of emotions	Relationship between empathizer and target	Situative context	Empathizer
Valence	Affective link and nurturance	Appraisal of the situation	Mood arousal
Intensity	Self-implication (e.g. jealousy, anger)	Display of multiple emotions	Personality, gender and age
Saliency	Familiarity and similarity		Emotional repertoire
Primary versus secondary emotions	Communicative intentions		Emotional regulation capacities

Tab. 2 Scheme of the model collecting the modulatory factors of empathy (adapted from) (De Vignemont & Singer, 2006).

De Vignemont and colleagues' model also proposes two different ways in which the modulation of empathy can occur.



Tab. 3 Schematic representation of the early and late appraisal model of empathy (adapted from) (De Vignemont & Singer, 2006).

According to the late appraisal route the empathic response is directly and automatically activated by the perception of an emotional cue. Then the default rule is that there is always an empathic response. However, the automatic reaction can be modulated in a second time. Information about the general and personal context is processed in parallel. The outcome of the contextual appraisal process leads to the modulation of the empathic response. Notably, this modulation can either be ruled by top-down inhibitory or excitatory processes or by horizontal competition between different motivational processes. This latter case could contribute to explain Singer et colleagues' data (Singer et al., 2006): different motivational systems might have competed while subjects were evaluating contest cues and the men's desire for revenge might have taken the upper hand of the inclination to empathize with someone feeling pain.

On the other hand, the early appraisal model postulates that the empathic response is not directly and automatically activated by the perception of an emotional cue. Rather, the emotional cue is evaluated in the context of external and internal information. Whether an empathic response is elicited depends on the outcome of the contextual appraisal process. Thus, in this latter model, the default

rule is that an empathic response is not automatically activated but an empathic response might be elicited as the outcome of the appraisal process

## **Empathy and social neuroscience**

In recent years a larger body of research has investigated possible modulatory affects of social factors on empathy for pain. Most of these studies focused on the relationship between the empathizer and the target and analyzed when and why our mind, our brain and our body decide not to share others' people pain.

Whether and how empathy is modulated by social factors has been a central topic in social and developmental psychology for many dacades. In 1975 Krebs (Krebs, 1975) investigated whether similarity, empathy and altruism are related. Subjects have been monitored Psychophysiological (SCR and HR) indexes have been measured while subjects were watching a person (a confederate) receiving reward, punishment or nothing at a gambling game. Furthermore, subjects were told that they had been paired with the performer on the basis of a computer analysis of their responses to the personality tests. Half of the subjects were informed that they had been paired with the performer because they were similar to him, and half were informed that they had been paired with the performer because they were different from him. According to the author, the similarity manipulation was designed to increase the disposition of similar subjects to empathize with the performer. It was predicted that subjects in the high-similar group would evidence the strongest psychophysiological reactions and congruent emotional responses when the performer would receive rewards and punishments but not when the performer was in the neutral condition. Data confirmed hypothesis by showing that subjects in the high-similar group evidenced the greatest increases in skin conductance ( $t = 2.01, p < .025$ , heart rate deceleration ( $t = 1.65, p = .06$ ), and heart rate acceleration ( $t = 1.30, p < .10$ ) (Fig. 9).

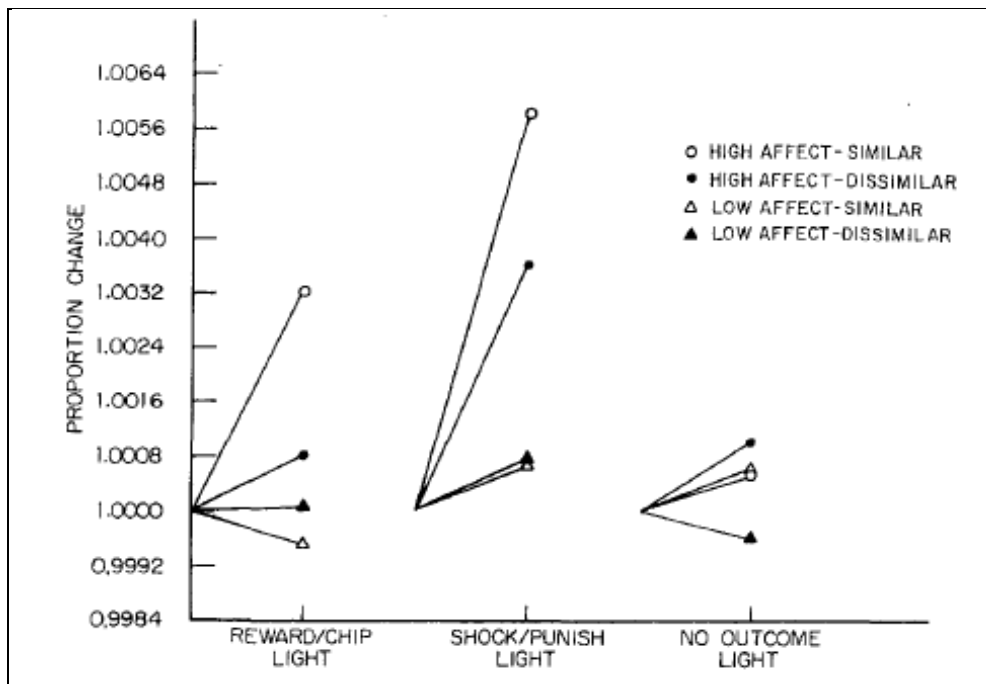


Fig. 9 Proportion of change in log skin conductance to lights (adapted from) (Krebs, 1975).

Furthermore, subjects in the high-similar group behaved most altruistically. They received an altruism score of 14.0, compared with scores of 9.6 for subjects in the high dissimilar group. Thus, according to the author, the major finding of the study was that subjects who experienced the strongest empathic reactions toward another were most willing to help him, even though it meant jeopardizing their own welfare.

Also neuropsychological data seem to show that empathy can be modulated by perception of contest cues.

In 2006 Singer and colleagues (Singer et al., 2006), with a fMRI study investigated whether empathy for pain is affected by affective link between the empathizer and a target person in pain. To address this question, the authors measured brain responses when individuals empathized with the pain of someone they liked or disliked. In order to modulate the fairness of the target person, they used an economic game model Prisoner's Dilemma game (PDG) that induces

liking or disliking for the two confederate actors, previously unknown to the experimental subjects, depending on whether they play fairly or unfairly. In the second part of the experiment they used functional magnetic resonance imaging (fMRI) to investigate whether the fairness of the confederate modulates empathic responses for pain. By using the same paradigm of a previous study (Singer et al., 2004), authors measured brain activations in the subjects both in pain/no pain self condition and in pain/no pain other condition. Data showed pain related empathic responses in both genders in bilateral AI extending into FI and brainstem when seeing an unfamiliar but likeable person in pain. To investigate whether empathic responses are modulated as a function of the perceived fairness of others, average activation have been calculated in bilateral FI for painful–non-painful stimulation when subjects observed either fair or unfair players in pain. Data showed a less empathic activity elicited by the knowledge that an unfair player was in pain. But notably, this empathic reaction is modulated by subjects' gender. In women this reduction in activity was very small, whereas in men the knowledge that an unfair player was in receipt of pain elicited no increase in empathic activity in FI. Statistical analysis revealed no significant difference for women when comparing painful vs no painful conditions for fair versus unfair players. On the other hand men showed significantly enhanced activation in bilateral FI when observing fair compared with unfair players in pain (Fig. 10). Furthermore, for men but not for women have been observed increases in brain activity in regions associated with reward processing (ventral striatum/nucleus accumbens and orbito-frontal cortex) when observing an unfair player receiving painful stimulation.

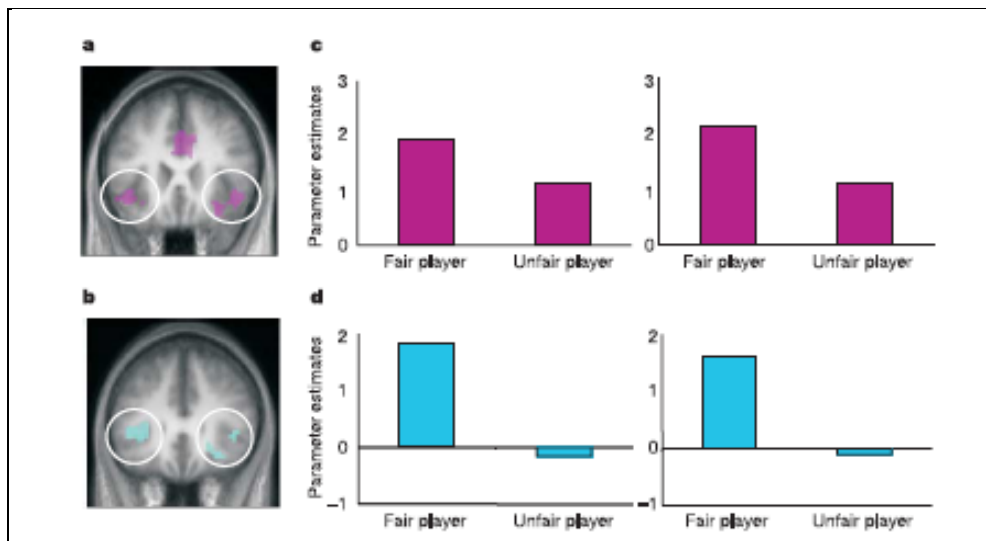


Fig. 10 Pain-sensitive activation networks to the sight of fair and un fair players in pain. a, b, Conjunction analysis between the contrasts pain–no pain in the context of self and the fair condition for women (pink, a) and men (blue, b). c, d, Average activation (parameter estimates) in peak voxels of left and right FI (left and right panels, respectively) for the painful–non painful trials in fair and un fair conditions for women (c) and men (d) (adapted from) (Singer et al., 2006).

This data provide neurobiological evidence on how fairness in social interactions shapes the nature of the affective link between people. This suggests that people like cooperating with fair opponents but also like punishing unfair opponents. This pattern of data is more prominent in men, maybe because the experimental design favored the modality of punishment related to physical threat. On the other hand, these findings could indicate a predominant role for males in the maintenance of justice and punishment of norm violation in human societies.

Taking together, Singer and colleagues' findings indicate that empathy can be clearly and deeply modulated by social factors like the affective link between the observer and the target.

In a similar vein, in 2009 Xu and colleagues (Xu, Zuo, Wang, & Han, 2009) studied whether empathy related brain activations is modulated by race based ingroup/outgroup categorization. They scanned Caucasian and Chinese participants using functional magnetic resonance imaging (fMRI) while they

watched video clips of Caucasian or Chinese faces receiving painful stimulation (needle penetration) or non-painful stimulation (cotton Q-tip touch). This experimental design allowed the authors to test empathic brain reactions and modulatory effect of target race as well as a possible three way interaction with subjects' race. Data showed that there was a significant interaction of Pain\*Group Membership, as ACC empathic responses were greater to racial in-group than outgroup members. However, the triple interaction of Pain\*Group Membership\*Ethnicity was not significant. According to the authors this reveals that there is a similar pattern of modulation of ACC empathic responses by racial group membership in Caucasian and Chinese participants (Fig. 11)



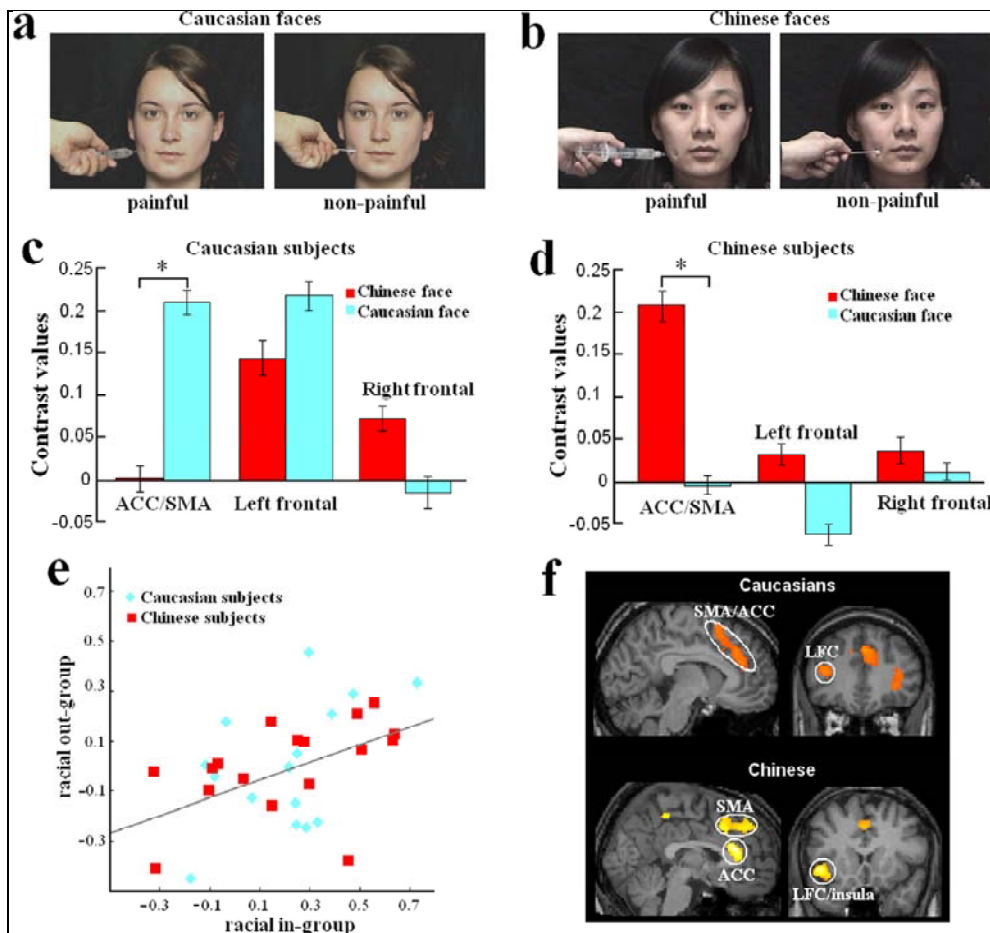


Fig. 11 A) Illustration of Caucasian faces receiving painful and non-painful stimuli. B) Illustration of Chinese faces receiving painful and non-painful stimuli. C) Contrast values of the parameter estimates of signal intensity in the ACC and the frontal cortex that differentiated painful and non-painful stimuli in Caucasians. D) Contrast values of the parameter estimates of signal intensity in the ACC and the frontal cortex that differentiated painful and non-painful stimuli in Chinese. E) Correlation between ACC empathic neural responses to racial in-group and out-group members. X and Y axes respectively indicate ACC empathic responses to racial in-group and racial out-group members indexed in contrast values of painful versus non-painful stimulation. F) Increased activations in the ACC and the frontal/insula cortex shown in whole-brain statistical parametric mapping analyses when participants perceived racial in-group faces. The upper figures show the results from Caucasian subjects and the lower figures show the results from Chinese subjects (adapted from) (Xu et al., 2009).

Taking together these findings constitutes a first and clear evidence of modulation of empathic neural responses by racial group membership, i.e., ACC empathic responses to perception of others in pain decreased remarkably when participants viewed faces of racial in-group members relative to racial out-group members. Furthermore this effect was comparable in Caucasian and Chinese

subjects and suggests that modulations of empathic neural responses by racial group membership are similar in different ethnic groups.

According to the authors, as race helps defining in-group/out-group members (Cosmides, Tooby, & Kurzban, 2003) on the base of previous studies showing that the ACC mainly contributes to the affective component of empathy (Singer et al., 2004), (Xu et al., 2009) data suggest that the own-race bias in ACC activity linked to empathy for pain may mediate enhanced sharing of feelings and emotions of ethnic in-group members. Thus it is likely that the own-race bias in empathy-related ACC activity observed here reflected unconscious affective response to racial in-group members.

An even more recent study in this very same field has been conducted by Avenanti and colleagues (Avenanti et al., 2010). The authors using transcranial magnetic stimulation explored sensorimotor empathic brain responses in black and white individuals. Specifically they sought to determine whether neurophysiological and autonomic indices of reactivity to others' pain are modulated by racial membership and racial bias. They analyzed a very basic form of interpersonal reactivity called sensorimotor contagion, which is indexed by an automatic reduction of the corticospinal excitability of onlookers who observe painful stimuli delivered to a stranger model. Using transcranial magnetic stimulation (TMS), the study aimed at exploring changes in excitability of corticospinal body representations in white-Caucasian (Italian) and black-African (born in Africa and living in Italy).

Subjects were asked to watch clips depicting (1) needles penetrating the right first dorsal interosseus (FDI) or (2) a Q-tip touching the very same hand muscle of stranger black or white models. Motor-evoked potentials (MEPs) to

single-pulse TMS of the left motor cortex were recorded from the observers' right FDI (target) and abductor digiti minimi (ADM, control) hand muscles.

Data showed that in the muscle x model ANOVA model on MEP difference the main effect of model is significant ( $F(1,34)=8.21, p=0.007$ ) and the model x muscle interaction is significant as well ( $F(1,34)=4.50, p=0.041$ ); specifically this was accounted for by the greater inhibition recorded from the FDI muscle during observation of ingroup rather than outgroup models ( $p = 0.002$ ) (Fig. 12).

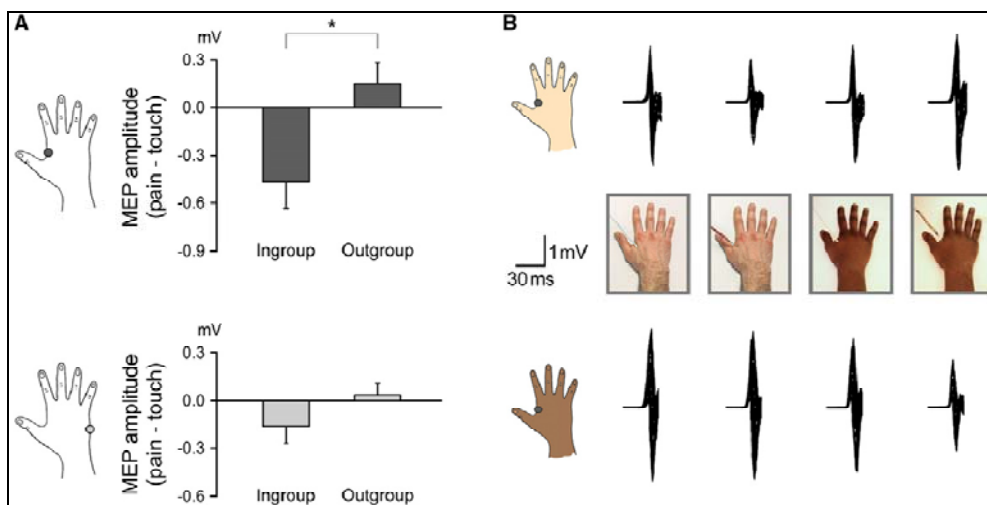


Fig. 12 (A) Mean motor-evoked potential (MEP) difference (pain - touch) recorded from the first dorsal interosseus (FDI) (dark gray) and the abductor digiti minimi (ADM) (light gray) muscles during the observation of stimuli applied to the ingroup and the outgroup models. The asterisk denotes significant post hoc comparison. Bars indicate standard error of the mean (SEM). Observing ingroup but not outgroup models' pain led to resonant inhibition of the FDI muscle that was stimulated in the models: one-sample t tests confirmed that MEP contrasts (pain - touch) for the FDI (target) muscle were significantly different from 0 for the ingroup ( $t(34) = 22.8, p = 0.007$ ) model, but not for the outgroup ( $p = 0.3$ ) model. For the ADM (control) muscle, MEP contrasts were not different from 0 for either model ( $P_s > 0.2$ ; see also Table S1). (B) Raw MEPs recorded from the FDI muscle in a white (top) and a black (bottom) representative subject (adapted from (Avenanti et al., 2010)).

These findings indicate that seeing pain in members of the same racial group induced a reduction of corticospinal excitability that was specific to the muscle that participants observed being penetrated. This means that watching stimuli supposedly painful to others induces a specific corticospinal inhibition,

suggesting the presence of a resonant activation of pain representations in the onlooker's sensorimotor system.

The subjects also completed a race IAT to assess the implicit racial bias White/Black races.

The IAT scores in white subjects indicated a implicit preferences for Whites (Italian) on Blacks (African Italian) ( $T(17) = 9.77, p < 0.0001$ ), indicating that they were quicker to associate concepts of good with the term "Italian" rather than with the term "African" and concepts of bad with the term "African" rather than with the term "Italian." A significant racial bias effect was found also in black (African) participants  $T(17) = 2.19, p = 0.043$ , indicating an implicit preference for the ingroup with the same pattern as for the Italian sample. The IAT scores crucially indicated that the ingroup-specific pain embodiment paralleled the implicit preference for ingroup members. Despite the lack of explicit racial bias in the sample, participants with higher implicit ingroup preference presented greater differences in the corticospinal reactivity to ingroup and outgroup models' pain ( $r = 20.46, p = 0.005$ ). This effect was similarly present in both white ( $r = 20.50, p = 0.033$ ) and black ( $r = 20.71, p = 0.002$ ) (Fig. 13).

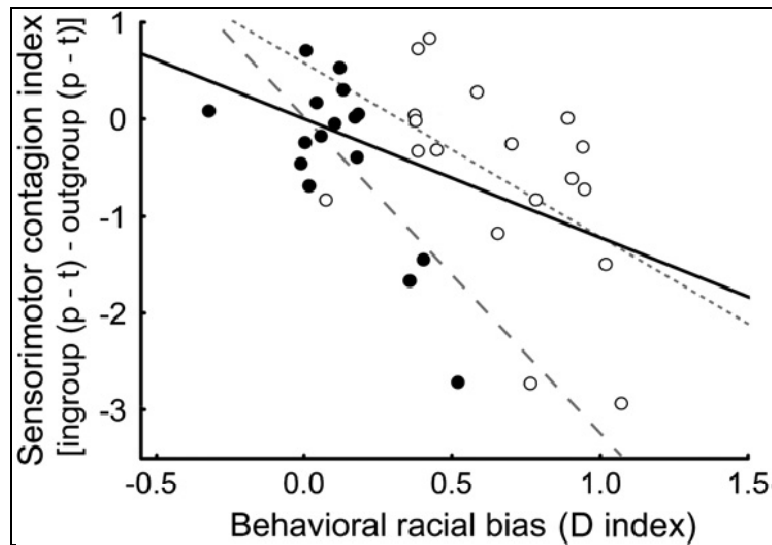


Fig. 13 Correlation Analysis between Behavioral and Neurophysiological Markers of Racial Bias. The regression lines indicate the correlation of the entire sample (thick line,  $r = 20.47$ ,  $p = 0.005$ ), the black subjects (stippled line,  $r=20.71$ ,  $p = 0.002$ ), and the white subjects (dotted line,  $r =20.50$ ,  $p = 0.033$ ), respectively. Negative correlations indicate greater sensorimotor response to ingroup relative to outgroup models' pain in those subjects who scored high on the race Implicit Association Test. White and black dots indicate whites and blacks (adapted from) (Avenanti et al., 2010).

These findings suggest a deep link between sensorimotor contagion and implicit race-related preferences. This link discloses social sensitivity in the human sensorimotor system and indicates that markers of social categorization can be found at basic sensorimotor levels of brain processing (Avenanti et al., 2010).

Authors also showed in a sharp way that the lack of empathy for the outgroup can not be accounted for by mere lack of similarity with the outgroup members.

They tested a subgroup of participants in two additional conditions in which pain or tactile stimuli were delivered to a violet-colored hand, i.e. a hand that does not define any groups socially defined.

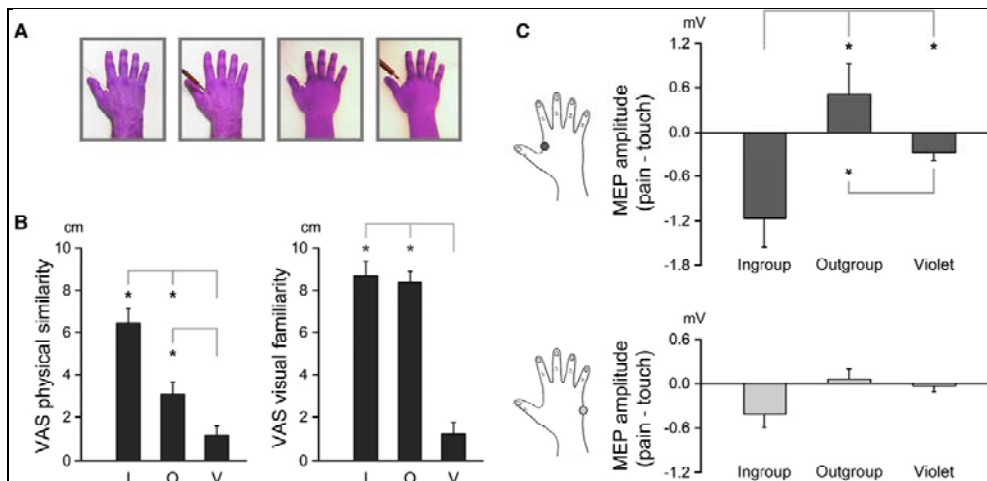


Fig. 14 Sensorimotor Contagion of Ingroup, Outgroup, and Unfamiliar Violet Models' Pain. (A) Examples of touch and needle in violet models, to which racial prejudices and stereotypes did not apply. (B) Subjective ratings (visual analog scale, VAS) of visual familiarity and of physical similarity of the observed hand with respect to the self hand. Subjective ratings indicate that the violet model (V) was judged as more unfamiliar and dissimilar than the ingroup (I) and the outgroup (O) models. Asterisks denote significant post hoc comparisons. See Supplemental Results for details concerning the statistical analysis of VAS data. Bars indicate SEM. (C) MEP difference (pain – touch) in the subgroup of onlookers tested during observation of ingroup, outgroup, and extremely unfamiliar violet models. Ingroup and outgroup data presented in this figure are a subset of the data from Figure 1. Asterisks denote significant post hoc comparisons. Bars indicate SEM. Sensorimotor contagion was found for both ingroup and violet models, but not for the outgroup model: one-sample t tests on MEP contrast (pain – touch) recorded in the target FDI muscle were significantly different from 0 for the ingroup ( $t(7) = 23.1, p = 0.02$ ) and the violet ( $t(7) = 22.3, p = 0.05$ ) models, but not for the outgroup model ( $t(7) = 1.2, p = 0.3$ ). For the ADM (control) muscle, MEP contrasts were not different from 0 for any of the models (all one-sample t tests:  $P_s > 0.1$ ) (adapted from) (Avenanti et al., 2010).

The muscle x model ANOVA on MEP differences showed a main effect of model ( $F(1,14)=7.05, p=0.008$ ) and a marginally significant model x muscle interaction ( $F(1,14) = 3.33, p = 0.065$ ). The inhibition found in the FDI muscle for the ingroup models' pain was greater than the inhibition found for the outgroup ( $p = 0.004$ ) and the violet ( $p = 0.02$ ) models' pain; and most importantly, the inhibition was greater for the violet than for the outgroup model's pain ( $p = 0.05$ ).

The outgroup model was explicitly rated as more familiar and similar to the self than the violet model. However, at a very basic level, a clear sensorimotor contagion was found for violet hand but not for the outgroup hand. Therefore, the

absence of pain resonant mapping cannot be explained by a reduction of observers' familiarity or by somatic similarity with outgroup members.

Data showed that the differential reactivity to ingroup and outgroup was predicted by racial bias suggesting that cultural conditioning (e.g., racial stereotyping), rather than biological or structural factors (e.g., somatic similarity), may shape embodied resonance with others. Thus, authors suggest that the basic reactivity of human beings implies empathy with the pain of stranger individuals. This reactivity may be maximal when the perceived similarity with the model is high (ingroup model) but is also present for very unfamiliar others if no stereotype can be applied to them (violet hand model). Crucially, race bias and stereotypes (outgroup model) may change interpersonal reactivity into a group specific lack of sensorimotor resonance (Avenanti et al., 2010).

Taking together these data show that our mind, brain and body do perceive social cues people communicate even if we are not aware of. Bindings with other people are affected by these cues and sharing of feelings and emotions of other people depends on how we perceive them. Who he/she is and who I think he/she is on the base of my knowledge is a key factor in determining how my body and mind react to his/her feelings.

Researches I conducted want to study why what I think of others is so important in determining my brain and body functioning and what are the key cues that play a central role in this kind of processes.

# **Racism and the empathy for pain on our skin**

## **Introduction**

Empathy is the ability to understand and vicariously share the feelings and thoughts of other people (De Vignemont & Singer, 2006). Empathic feelings are fundamental for humans in social and interpersonal life because they enable human beings to tune their mental states to their social environment as well as to understand others' intentions, actions, and behaviors. One of the main sources of empathic feelings is the pain experienced by other human beings, and empathy for others' pain, in turn, regulates behavior among individuals and social groups. Feeling other people's suffering triggers prosocial behavior (Batson, Chang, Orr, & Rowland, 2002), promotes helping and encourages cooperation (C. D. Batson et al., 1997). Conversely, lack of empathy for the pain of other human beings may lead to violence, abuse, and deterioration of interpersonal and intergroup relationships (Batson et al., 2002).

Empathic reactions to pain involve different layers of cognitive processing, with a predominant role played by automatic and implicit processes. Recent neuropsychological findings (Singer et al., 2004) have documented specific neuropsychological activations of the affective but not sensory components of the brain (the pain matrix in particular), leading to fast and automatic responses to the pain of others. Similarly, the vision of a needle penetrating the hand reduces the muscular motor response in the observer compatible with the locus of injection in the target person (Avenanti, Buetti, Galati, & Aglioti, 2005). Immediate empathic reactions, however, depend on the characteristics of the target person of the pain. When the target person has behaved uncooperatively, for instance, empathic reactions decrease (Singer et al., 2006). Furthermore, when the target person is



perceived as more similar to the observer, empathic reactions are stronger and trigger more altruistic behavior (Krebs, 1975). The importance of the characteristics of the sufferer is evident even in a phylogenetic perspective: The physiological empathic responses of human observers observing victimized animals are stronger when the animals are phylogenetically closer to humans (Rae Westbury & Neumann, 2008). Thus, an important question in understanding the role of empathy in regulating interpersonal relationships regards the possible differences in empathic reactions due to the important social characteristics of the person in pain.

Race is a prominent social factor characterizing modern multi-cultural societies. If empathic reactions differ depending on the race of the target person, helping behavior and cooperation among members of different racial groups can be undermined and disrupted.

With these experiments, I aimed at providing experimental evidence that automatic, physiological reactions to other people's pain strongly depends on the race of the person in pain, such that pain received by members of other racial groups elicits a much weaker reaction compared with the pain suffered by members of the same group. In two experiments I tested whether the reaction to pain of Caucasian (Italian) observers was influenced by the race (Caucasian, Asian or African) of the person in pain. In the second study we replicate this finding and show that the moderation of empathy is correlated with the individual implicit racial biases.

Empathic reactions were measured via Skin Conductance responses (SCR) (Dale et al., 2008) to observed video stimuli showing human subjects experiencing either harmless or painful somatosensory stimuli.

The third-person exposure to pain activates in the brain network called “pain matrix” (Derbyshire, 2000; Peyron et al., 1999; Peyron, Laurent, & García-Larrea, 2000) which comprehends the ACC. Activity in Anterior Cingulate cortex is known to trigger variation in the Skin Conductance (Dale et al., 2008), i.e. skin conductance increases as a physiologic and autonomic response to someone else’s pain (Krebs, 1975; Levenson & Ruef, 1992; Morrison et al., 2004; Rae Westbury & Neumann, 2008). The triggering role of the ACC on skin conductance is likely to have been selected to facilitate coping and adaptive responses (Devinsky, Morrell, & Vogt, 1995; Ledowski et al., 2006). Skin conductance is considered one of the most reliable predictors of accurately assessed negative emotions in others but it is associated with emotional responses rather than being specifically paired with pain recognition (Levenson & Ruef, 1992). This neural and physiological activity is critical in empathy for pain such that individuals with lesions in the ACC or impaired autonomic control have diminished empathic abilities (Chauhan, Mathias, & Critchley, 2008; Fujiwara et al., 2008).

## **Experiment 1: Racism and indifference to the pain of other people.**

### **Method**

<sup>1</sup>Each video started with a frame depicting a face of a female or a male actor holding a neutral expression. Subsequently the camera zoomed in on the actor’s hand which was touched by the experimenter alternatively by an eraser (harmless stimulus) or by a needle (painful stimulus). Twelve video clips were prepared using six different actors, one male and one female actor for the Caucasian, African and Asian races.

---

<sup>1</sup> Findings described in experiments 1 and 2 have been submitted to “Frontiers in Perception science”.

In both experiments<sup>2</sup>, participants sat in front of a computer monitor (Acer aspire 1360, monitor 15.4" TFT LCD) where the experimental stimuli were displayed. Prior to stimulus presentation, two electrodes were applied on the forefinger and ring-finger of participant's left hand in order to record the SCR. Participants were asked to relax, and carefully watch the stimuli presented on the monitor. During SCR recording, participants were listening to white noise with headphones in order to cover external auditory stimuli. The videos order was completely randomized. The experimenter, blind to stimuli presentation, started each video after visually checking that the online SCR was returned to a baseline level, in order to avoid response overlaps to consecutive stimuli. Following this procedure the inter stimulus interval was 15 seconds (range 10 to 20) across all participants. All participants gave their consents to physiological recording and display of videos prior to the experiments. After the experiments ended, participants were fully debriefed regarding the nature of the stimuli and aim of the study.

## **Measures**

SCR was measured while the observers viewed the video stimuli. The difference between the SCR subsequent to a painful stimulus and the SCR subsequent to a harmless stimulus was taken as a measure of empathy for pain (hereinafter EI, Empathic Index).

Physiological data collection was performed using The UFI model 2701 BioDerm(R) Skin Conductance meter. It is a stand-alone instrument which measures Skin Conductance Level (SCL) and Skin Conductance Response (SCR). Skin conductance is measured using an Ag-AgCl electrode pair with the constant

---

<sup>2</sup> Both of the experiments have been conducted in laboratories of the Department of Psychology of the Milano-Bicocca University in Milan.

voltage (0.5 volts) method. The SCR were recorded as the phase component of the skin conductance activity, with a 10Hz rate. Microsiemens ( $\mu\text{s}$ ) are the measurement units. For all the participants, the SCR recordings were synchronized with the first video frame presented. The average response within a time window of 6.5 seconds post-stimulus was used as the observed variable for the analysis<sup>3</sup>. This interval was chosen based on the relevant literature as the most appropriate and included the whole variation of SC following the stimuli. The SCR data analysis was performed using the SAS GLM procedure. Unless otherwise specified, all the results discussed were obtained with a least squares repeated measures ANOVA. The estimated means for the stimulus X race interaction at different levels (1 SD below and 1 SD above the mean) of the continuous IAT variable were obtained with an equivalent model estimated with the PROC MIXED SAS procedure.

## **Subjects**

Ninety students of Milano-Bicocca University have been recruited. Three participants were excluded for problems in data saving; five participants were excluded because of uncooperative behavior during the experiment; seventeen participants were excluded due to technical problems during the experiment. Out of the 65 remaining participants, four were excluded as outliers: Skin Conductance Response (SCR) scores exceeding 2sd from the overall average. A total of 61 participants (29 female) were therefore included in the analyses.

Participants were subjected to a 2 (stimuli: harmless and painful) X 3 (races of the target person: African, Caucasian, Asian) X 2 (blocks: 1<sup>st</sup> and 2<sup>nd</sup>)

---

<sup>3</sup> This type of analysis is consistent with technical manual published by the producer of the UFI Bioderm System (<http://www.ufiservingscience.com/>).

experimental block) X 2 (target gender: male and female) repeated-measure factorial design.

## Results

Participants showed an overall significant EI: reactions to painful stimuli were significantly greater than reactions to harmless stimuli [ $F(1,59) = 40.85, P < 0.001$ ].

Crucially, the race of the actor experiencing the painful stimulus significantly moderated the EI [ $F(2,118) = 3.6, P = 0.03$ ]. Although experimental participants showed a significant EI for Caucasians [ $F(1,59) = 29.57, P < 0.001, \eta^2=.333$ ], for Africans [ $F(1,59) = 7.52, P = 0.008, \eta^2=.113$ ] and for Asian images [ $F(1,59) = 16.99, P < 0.001, \eta^2=.223$ ], the empathic reaction for the Caucasians was significantly greater than that for the Africans [ $F(1,59) = 7.87, P = 0.006, \eta^2=.0,117$ ] (Fig. 15).

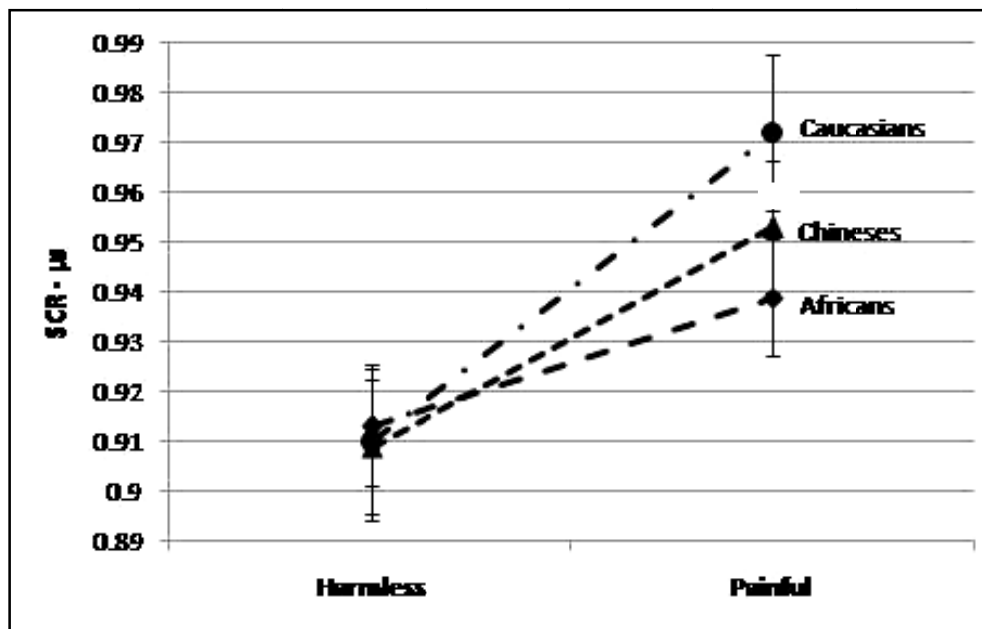


Fig. 15 Experiment 1: mean SCR and standard errors as a function of stimulus type and actor's race. Responses to pain were always greater than those to the harmless stimuli. Critically, there was no racial effect on the reaction to the harmless stimuli [ $F(2,118) = 0.09, P = .91$ ], a significant moderating effect of target person race was found on the reaction to painful stimuli [ $F(2,118) =$

5.09,  $P = .007$ ): Reactions to Caucasians painful stimuli were significantly greater than for Africans [ $t(118) = 2.91$ ,  $P = .004$ ] but not than for Asian targets [ $t(118) = 1.72$ ,  $P = .08$ ].

## **Experiment 2: implicit attitudes and empathy for pain**

### **Measures**

The second experiment aimed at linking the empathic racial bias with the implicit racial prejudice. In addition to using the same paradigm used in Experiment 1, the experiment required participants to complete a race (Caucasians/Africans) Implicit Association Test and a Trait Empathy Scale (Mehrabian & Epstein, 1972). Specifically, we assessed to what extent individual differences in the implicit racial prejudice correlate with the difference between the empathic index for Caucasians with respect to EI for Africans.

The Implicit Association Test (IAT) provides access to deep cognitive domain that is not reached by self-report measures (Greenwald, McGhee, & Schwartz, 1998). In the version used in this study, it provides a measure of implicit differential evaluation of Caucasian and African races. The IAT is based on participants' reaction times on a computer-based categorization task. The IAT assesses the association between two classes of stimuli by measuring differences in the response speed that participants show in the same task with exemplars from two categories. The task we used rates the association strength between positive and negative concepts with Caucasian and African races. On each trial of the race IAT we used, participants categorized a stimulus from one of four the categories: a photo of a Caucasian man, a photo a African man, a positive word (Joy, love, peace, wonderful pleasure, friend), or a negative word (agony, terrible, awful, bad, evil, war). In one block of trials, positive words required the same behavioral response as photos of Caucasian men. In another block of trials, positive words required the same response as photos of African men. IAT data were coded in the

direction of association between positive words and Caucasian targets, i.e. as the difference in mean response latency to trials in positive word-African targets block minus trials in the positive words-Caucasian targets block. Then, higher scores reflect strong associations between positive concepts and Caucasian race as well as strong association between negative concepts and African race. The 4<sup>th</sup> and 7<sup>th</sup> blocks consisted of 40 trials, all of the other blocks consisted of 20 trials. In the data analysis, individual IAT scores were used in the mixed model as a continuous independent variable. The simple slopes analysis was conducted to estimate the experimental effects at specific values of the IAT score. This was obtained by centering the IAT score to one standard deviation above the mean (and subsequently at one standard deviation below) before entering the IAT score in the model (6).

## **Subjects**

This experiment included sixty students of Milano-Bicocca University. Two participants were excluded for problems in data saving; five participants were excluded due to technical problems during the experiment. Out of these 53 participants, six were excluded as outliers (SCR scores exceeding 2sd from overall average). A total of 47 participants (24 female) were therefore included in the analyses.

## **Results**

Results replicated the overall stronger reaction to painful than to harmless stimuli [ $F(1,45) = 36.63, P < 0.001$ ]. Target race significantly moderated the EI [ $F(2,90) = 4.26, P = .01$ ] (Figure S1). The EI was significant for Caucasian [ $F(1,45) = 23.85, P < .0001, \eta^2 = .346$ ] and Asian [ $F(1,45) = 13.9, P = .0005, \eta^2 = .225$ ] but not for African actors [ $F(1,45) = 1.36, P = .24, \eta^2 = .029$ ].

Furthermore, the EI was statistically lower for Africans than for Caucasians [ $F(1,45) = 6.64, P = .01$ ] targets (Fig. 16)

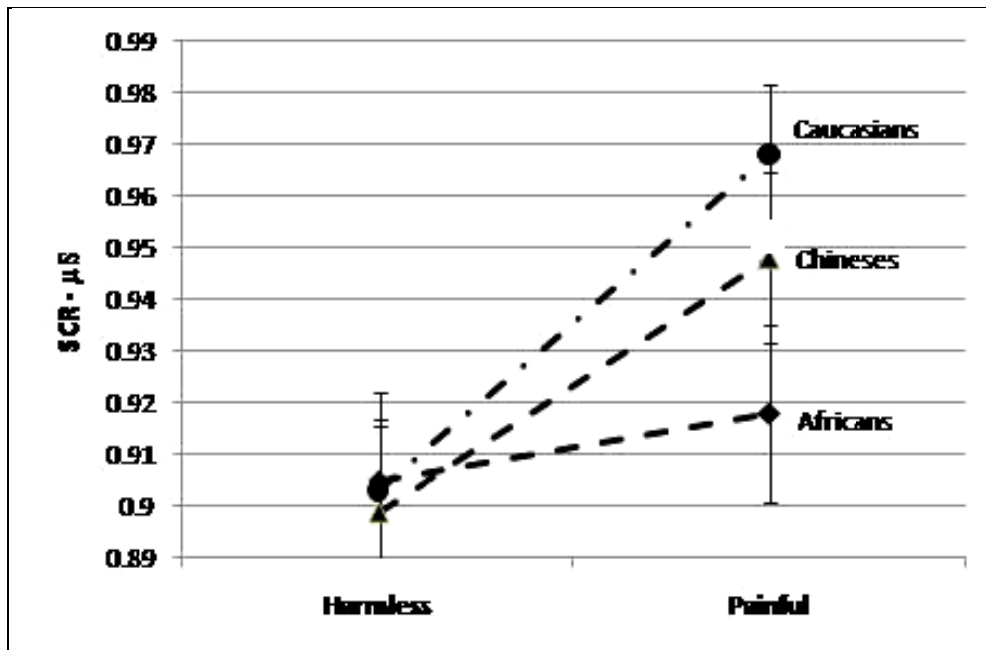


Fig. 16 Experiment 2: SCR means as a function of stimulus type and the race of the person in pain. As for experiment 1, we observed no race differences in reactions for harmless stimuli [ $F(2,90)=0.1, p=.9$ ] and a race significant effect on the reaction for other people's pain [ $F(2,90)=7.55, p<.001$ ].

The SC responses and the IAT scores were analyzed together. In the IAT used, greater scores indicate faster associations of “Caucasians” with positive concepts and “Africans” with negative concepts, thus a stronger racial bias against “Africans”. The IAT scores were included in an ANOVA comprising the SCR values elicited only by Africans and Caucasians stimuli (the two races included in the IAT). As typically found with the race IAT, Caucasian observers more strongly associated negative stereotypes with Africans than with Caucasians, [ $F(1,46) = 34.45, P < 0.001$ ]. Most importantly, the strength of the implicit race bias correlates with the reduced empathy for Africans' pain. We found that the IAT scores of the observers significantly predict the moderating effect of race on the reaction for pain [ $F(1,43) = 4.52, P = .03$ ]. Simple slope analysis (Aiken &



West, 1991) revealed that the greater the participant racial bias, the greater the difference between the empathic responses towards Caucasians with respect to Africans (Fig. 17). Data show that participants with low race bias (1 SD below sample average) are not affected by the race moderating effect on the empathic responses to actors' pain. Participants EI is significantly greater than zero [ $F(1,45) = 5.22, P = .02$ ] but it's not moderated by the race of the person in pain [ $F(1,45) = 0.14, P = .70$ ] and there are no differences in the overall reactions for Caucasians and Africans [ $F(1,45) = 1.28, P = .25$ ]. On the other hand, data show that for subjects with an high race bias (IAT score 1SD above the sample average) EI is significant greater than zero [ $F(1,45) = 14.52, P = .0001$ ] and the race of the person in pain significantly moderates the empathic reactions [ $F(1,45) = 13.29, P = .0003$ ].

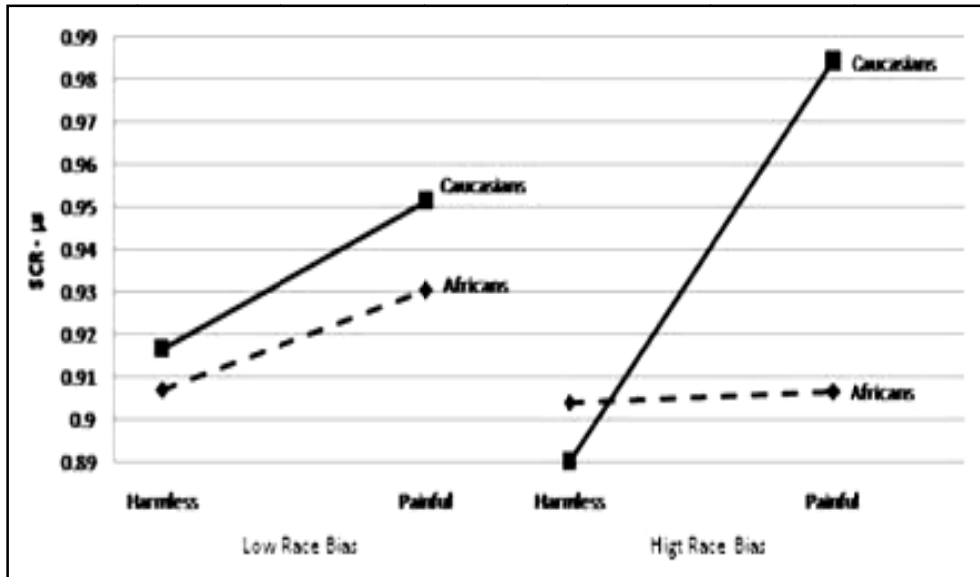


Fig. 17 Experiment 2: Simple slope analysis. Figure shows how the EI is significantly moderated by IAT subjects scores. Subjects with a strong negative bias for Blacks showed a greater EI for Whites than for Blacks. Subjects with a low race bias did not show any differences in the empathy for Whites' pain and for Blacks' pain.

The BEES empathy scale completed by participants had no significant effects on the SCR Empathic Index for any of 3 races we tested. Even though the BEES seems to account for the empathy related brain areas activation (Singer et al., 2004), this empathy scale seems to have failed in prediction of SCR empathy related activations (Rae Westbury & Neumann, 2008).

As hypothesized, the SCR responses significantly varied during the time for the painful stimuli, but not for the harmless stimuli. Specifically, the reactions to painful stimuli significantly reduced over time [experiment 1,  $F(1,59) = 44.58$ ,  $P < .0001$  and experiment 2,  $F(1,45) = 8.08$ ,  $P = .006$ ] and the reactions to harmless stimuli were constant during the experiments [experiment 1,  $F(1,59) = 0.15$ ,  $P = .70$  and experiment 2,  $F(1,45) = 1.67$ ,  $P = .20$ ]. These data suggest that participants' stimuli perception were reliable and precise during the entire experiment.

In order to rule out possible alternative explanations, we performed the analysis of SCR base-line values immediately before participants' empathic reactions. The mean of SCR values during the 600ms pre stimulus was calculated. As expected, in experiment 1 and 2 the full model revealed no relevant effects for all the experimental factors on the pre-stimulus SCR. These findings rule out the possibility that the observed responses associated with the painful and harmless stimuli were due to stochastic effects prior to stimulus presentation.

## **Experiments 1 and 2: general discussion**

Taken together our findings demonstrate a clear pattern of responses to pain: the extent to which Caucasian observers share the pain experience of other people is affected by the race of the person in pain (Fig. 18). Before the stimulus onset, the SCR values show stochastic variations. After observing a painful stimulus

administered to the target person, participants' SCR values increase more for Caucasian targets than for target people of the other races, and the least for African targets.

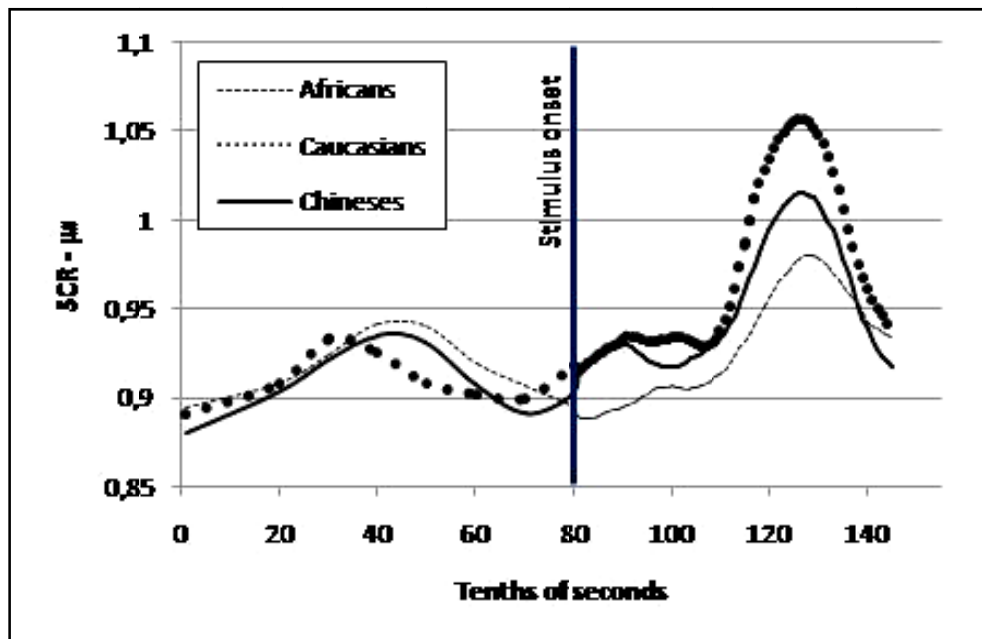


Fig. 18 Experiment 1 and 2, mean values. SCR means for painful stimuli as a function of time and actors' race.

This race moderation pattern was not present in the reactions to harmless stimuli (Fig. 19). During the video stimuli perception, before and after the stimulus onset, participants' responses are not affected by the race of the target people.

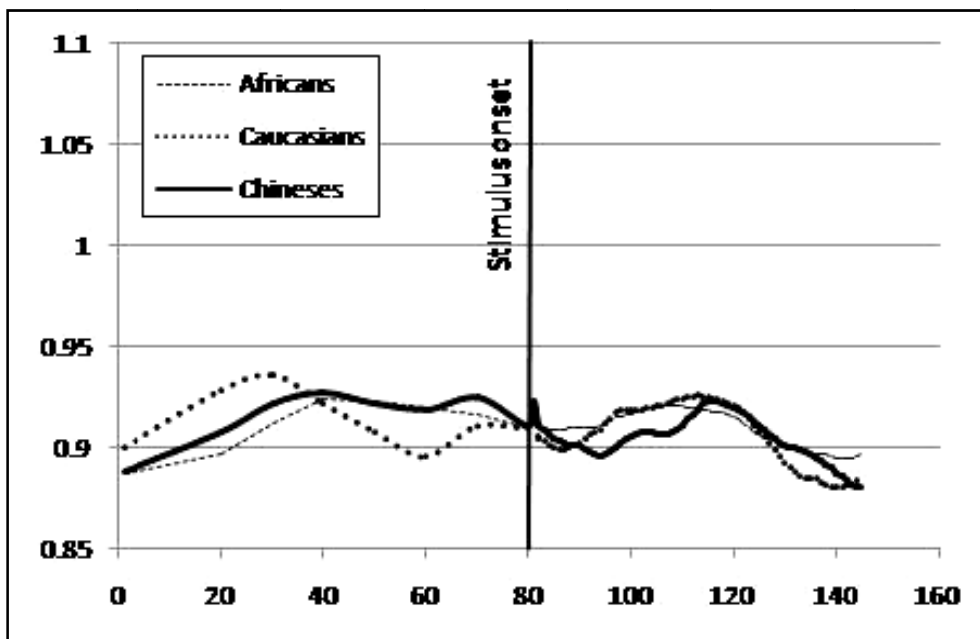


Fig. 19 Experiment 1 and 2, mean values: SCR to harmless stimuli as a function of time and race of target people.

Our data suggest that implicit and uncontrolled cognitive mechanisms lead Caucasians to reduce the automatic sharing of pain experience with African conspecifics at an automatic, early level of stimulus processing. These data concur with studies demonstrating deep connections between implicit social evaluations and neurological responses of the central nervous system. Research in the field of social neuroscience show that in white participants the strength of amygdale activation to unfamiliar black vs. white faces is correlated with implicit but not explicit race bias measures (Phelps et al., 2000) and is related to different level of PFC activity (Cunningham et al., 2004).

The correlation between the empathic reactions as measured with the SCR and the IAT scores rules out alternative explanations of the effect based on some low-level perceptual features of the video-stimuli presented. In fact, perceptual artifacts should have a general reduction of response to African actors for both harmless and painful stimuli, whereas the moderation due to race is specifically

found for painful stimuli. As our data did not show any gender effect, mere similarity between actors and observers could not account for our results.

Interestingly, these data do not support the outgroup antipathy hypothesis (Brown, Bradley, & Lang, 2006) as they do not indicate increased affective reactions to stimuli of outgroup members in general. Moreover, the ingroup empathy hypothesis (Brown et al., 2006) does not seem to account for our effects either. A mere ingroup-outgroup categorization should lead to a significant reduction of empathy for Africans as well as Asian actors. Instead, the EI to Africans was lower than the one for Caucasians and Asian targets in both experiments. Furthermore, although one of the most pervasive categorizations in human society is gender, we never observed an interaction between the subject's gender and stimulus gender in the empathic responses (in both of the experiments 1 and 2,  $P_s > .63$ ).

These data support the idea that racial groups different from the perceiver could elicit a weaker sense of familiarity than a more similar conspecific. *Dehumanization Theory* (Fiske, Cuddy, Glick, & Xu, 2002) posits that some specific social groups are perceived as less human as they activate non exclusively social emotions. Phenotypically distant outgroups may even elicit different patterns of brain activation, with particular regard to the medial PFC (Harris & Fiske, 2006), a region of cortex implicated in social cognition. This region responds to faces of people belonging to all social groups except extreme outgroups who activate, instead, a pattern consistent with disgust. Different degrees of dehumanization may therefore account for our findings that Caucasians could perceive Asians and Africans on different levels of humanization.

Findings regarding a differential reaction to animals in pain (Rae Westbury & Neumann, 2008) suggest that empathic feelings in humans are moderated by

the perceived phylogenetic similarity between the observer and the suffering animal. In a similar vein, data gathered in the 2 studies, seem to indicate that the closer the phenotypic aspect of the actor and the observer, the stronger the psychophysiological empathic response to pain.

My data suggest that the attitude towards other races may involve not only the overt self-report of the observer concerning attitudes about race but also their deep automatic and physiological reactions. These differential reactions may be elicited even at a very basic level, such as the reaction to physical pain of others. Such a fundamental racial differentiation, in turn, may bias complex activities and judgments over and beyond human consciousness. A precise assessment of other people's pain, in fact, is a necessary skill in many human activities, from medical decisions, rescue operations, police intervention, policy making and, in extreme circumstances, use of physical force and punishment. When all these activities involve people perceived as belonging to different races, a racial bias may hinder pain assessment with detrimental effects on individuals, groups, and their peaceful relationships.

## **Experiment 3: race and pain perception, an experimental study**

### **Introduction**

On the base of previous experiments, I conducted a third experiment<sup>4</sup> to further explore the role of race in determining empathy reactions to others' people pain.

By using two different samples (an African-American sample and a Caucasian American sample) I used the same paradigm as the one used in experiment one and two. With the new experiment I aimed at testing whether the reduction of empathy for Blacks people pain showed by Italian Caucasian subjects is due to the specificity of the target group or whether is a more general social phenomenon. In other words I wanted to understand whether lack of empathy for a social outgroup is due to social categorization mechanisms or whether it's due to specific, culturally determined cues of one social group.

### **Method**

Forty new videos have been made. Each video lasts for 8 seconds and portrays a Black or White target in a painful or harmless condition. In the painful condition, a target is injected on the back of the hand by a needle; in the harmless condition, the target is tapped on the hand by an eraser. The needle is attached to a syringe and it seems it penetrates the hand's skin. Actually a fake syringe has been used and no pain as been inflicted to the actors. The scene in the painful condition is intended to provide the illusion of pain. Targets include five Black males, five Black females, five White males, and five White females. Each video begins by presenting a target from the waist up, sitting at a table, and looking

---

<sup>4</sup> I conducted this experiment while I was working in the Implicit Social Cognition Laboratory led by prof. Mahzarin Banaji. The ISCL is part of the Department of Psychology at Havard University (MA, USA).

directly at the camera with a neutral expression. The camera then focuses in on the target's hand and lower chest. A gloved hand holding a syringe or eraser enters the frame from the upper right corner of the screen. The eraser or needle is pressed into the target's skin, after which the hand and stimulus exit the frame.

Each subjects viewed 40 videos corresponding to 8 experimental conditions: 2 stimuli (needle or syringe), x 2 races (Black or White actor), x 2 genders (male or female actor). For each experimental condition 5 videos have been made with 5 different actors.

Specifically, upon arrival, the study procedures are explained to the subject and any questions are answered. Informed consent is obtained once all questions or concerns are addressed. The subject sits in front of a computer screen and two electrodes are attached in bipolar placement to the palmar surface of the index and middle finger of the subject's right hand. The Biopack system (<http://www.biopac.com/>) collected and stored the SCR data. Before the videos begin, subjects are presented with instructions on a computer screen. They are told that they are about to see a series of videos during which either a painful or harmless object will be used on various individuals. Still pictures are shown of a needle and an eraser. The subjects are asked to remain as still as possible. After the videos are shown, subjects either complete the race version of the IAT followed by the IRI and BEES questionnaires or vice versa. At the completion of all the experimental measures, subjects are debriefed and financially compensated for their time.

SCR data have been analyzed with a 2 actor races x 2 stimuli x 2 actor genders x 2 subjects genders mixed model ANOVA.

The dependent variable is calculated starting from the frame in which the stimulus touched the actors' hand (5,5 seconds after videos onset) and considering



0,5 seconds for the response latency and a window of interest of 4,5 seconds. The dependent variable is the difference between the maximum SC value in the window of interest and the SC value at the very beginning of that window.

## **Measures**

During the experiment SCR has been measured on line while the subjects were watching the video-stimuli. After the experiment, subject completed the Interpersonal Reactivity Index (IRI) (Davis, 1980), the Balanced Emotional Empathy Scale (BEES) (Mehrabian & Epstein, 1972), and the race version of the Implicit Association Test (IAT) (Greenwald et al., 1998).

## **Subjects**

To test my hypothesis I tested 56 students (volunteers) located in Boston and Cambridge, (MA, USA). 16 participants have been excluded from the final sample: due to technical problems insufficient data were collected for 12 participants; 4 participants provided unusable data due to noncompliance. The final study sample (N = 40) includes 17 Black (4 female) participants who identified themselves as African/African American; and 23 White (11 female) participants who identified themselves as European American. All subjects are between the ages of 16 and 29. These participants were recruited for the study through an online pooling system used for studies conducted by the Department of Psychology at Harvard University.

## **Results**

Fig. 20 and Fig. 21 show the time windows used to calculate the DV: with a sampling rate of 25hz (25 SC values for second) figures indicate that the stimulus onset (the frame in which the stimulus touched the actors' hand) occurs 5,5 secs

after the beginning of the video. The latency window lasts for 0,5 sec and the window of interest lasts for 4,5 secs.

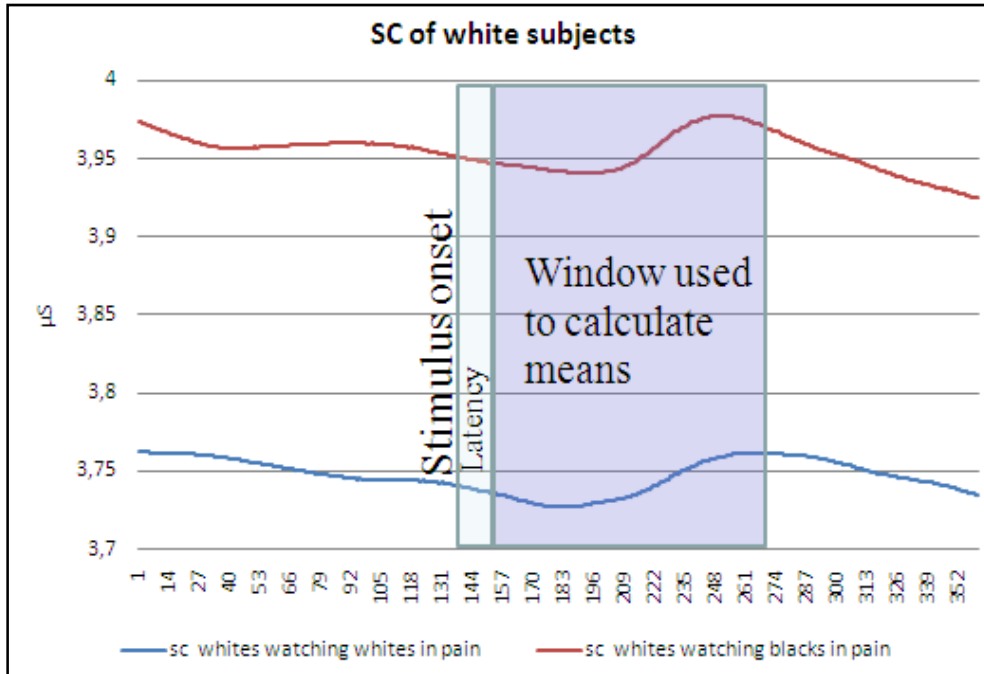


Fig. 20 Data show that in the painful condition Whites had a lower SC while were watching at Whites than when they were watching at Blacks.

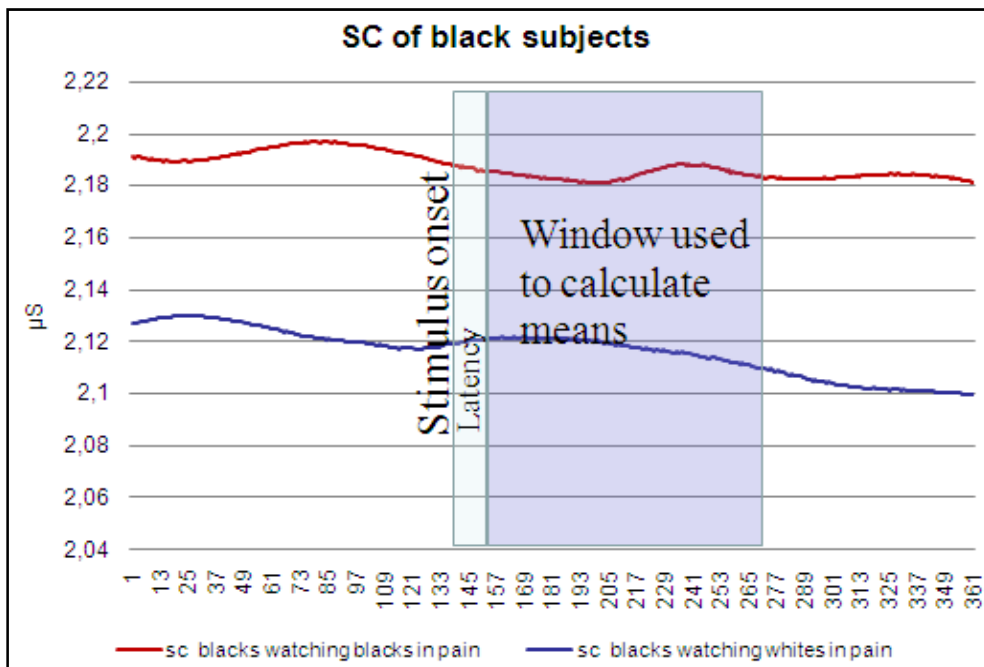


Fig. 21 Data show that in the painful condition Blacks had a lower SC while were watching at Whites than when they were watching at Blacks.

If we look at the painful condition, data indicate there is a subjects' race main effect  $F(1,38)=7,90$ ,  $p<0,001$  meaning that black subjects had a significantly lower SC level. A significant actors' race main effect emerged  $F(1,38)=9,88$ ,  $p<0,01$  meaning that black actors triggered a significantly higher SC level in the sample. But the interaction actors' race x subjects' race is not significant  $F(1,38)=1,27$ ,  $p=0,26$ . This means that black subjects had a higher SC level than Whites during all of the painful videos independently from the race of the actor they were looking at. At the same time data indicate that black actors triggered more SC to all of the subjects, both Blacks and Whites.

The IAT scores<sup>5</sup> tends to be significantly higher in white sample than in black sample  $T(34)=-1.97$ ,  $p=0.056$ , (whites mean=0.49, black mean=0.22). Data did not show significant differences between Whites and Blacks in the BEES scale ( $p>0.05$ ) as well as in perspective-taking, fantasy and personal distress subscales in the IRI scale ( $P_s>0.05$ ). Empathic concern scores were significantly higher in Whites  $T(38)=2.65$ ,  $p<0.05$ .

To test the experimental hypothesis I first analyzed separately White subjects' and Black subjects' data. Whites showed a significant main empathy effect  $F=(1,21)=10.25$ ,  $p<0.001$ , but the empathy reaction was not affected by the actors' race  $F=(1,21)=0.18$ ,  $p=0.6$ .

As shown in Fig. 22 whites reacted equally for the pain of Blacks as for the pain of Whites.

---

<sup>5</sup> For technical reasons we lost IAT data of 4 subjects, then we can include in this analysis 20 whites subjects and 16 black subjects.

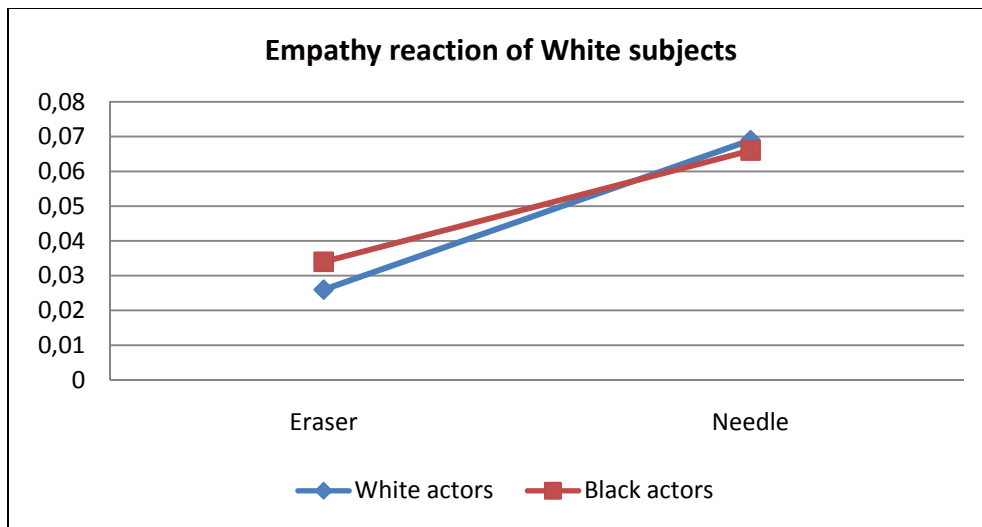


Fig. 22 Empathy reaction of White subjects for White and Black actors.

No other effects were significant. By considering in the model the Implicit Association Test scores I aimed at testing whether the implicit prejudice moderates the empathy reaction for the two different groups.

The IAT main effect was not significant  $F(1,16)=0,14$ ,  $p=0,7$  and, as predicted, it indicates that the IAT scores did not predict the SC values of the subjects.

Critically the three way interaction stimulus x actors' race x IAT is significant  $F(1,16)=4.78$ ,  $p=0.04$  and indicates that subjects' implicit race bias affects the different empathic reactions for Whites and Blacks.

In order to study the effect of the IAT on the empathy for pain, I estimated the stimulus x actors' race interaction for subjects with low IAT scores (scores under the sample mean) and for subjects with high IAT scores (scores above the sample mean). For subjects with a weak implicit racial bias (low IAT scores) the stimulus x actors' race interaction was not significant  $F(3,16)=1.95$ ,  $p=0.16$ . Fig. 23 shows that even though it's not statistically significant, data are not consistent with predictions. Whites with low IAT scores showed a slightly higher empathy for Whites than for Blacks.

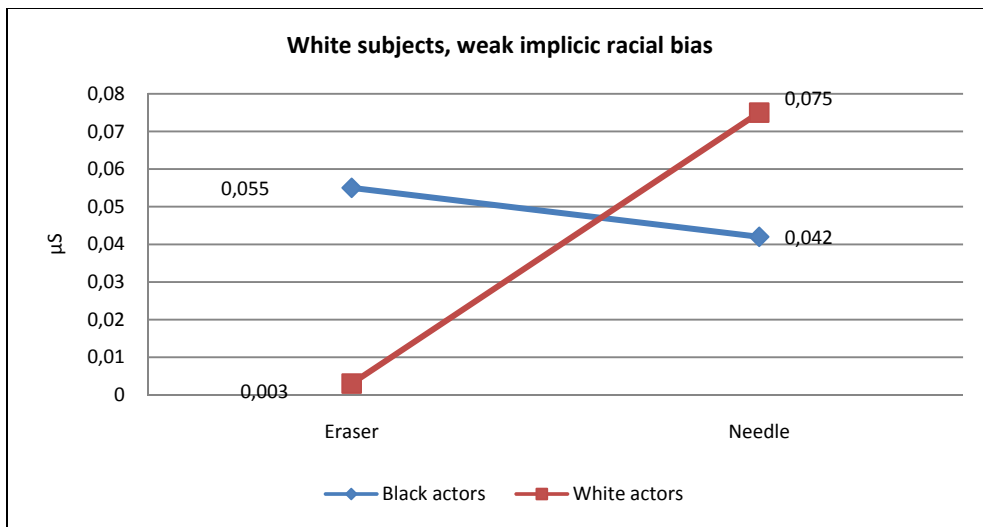


Fig. 23 Data show the empathic reactions of Whites with low IAT scores for White and Black actors.

For subjects with a strong implicit racial bias (high IAT scores) the stimulus x actors' race interaction was not significant  $F(3,16)=2.49$ ,  $p=0.10$ . Fig. 24 shows that even though it's not statistically significant, data show an unexpected trend. White subjects with strong IAT implicit negative bias for Blacks showed a slightly higher empathy for Blacks than for Whites.

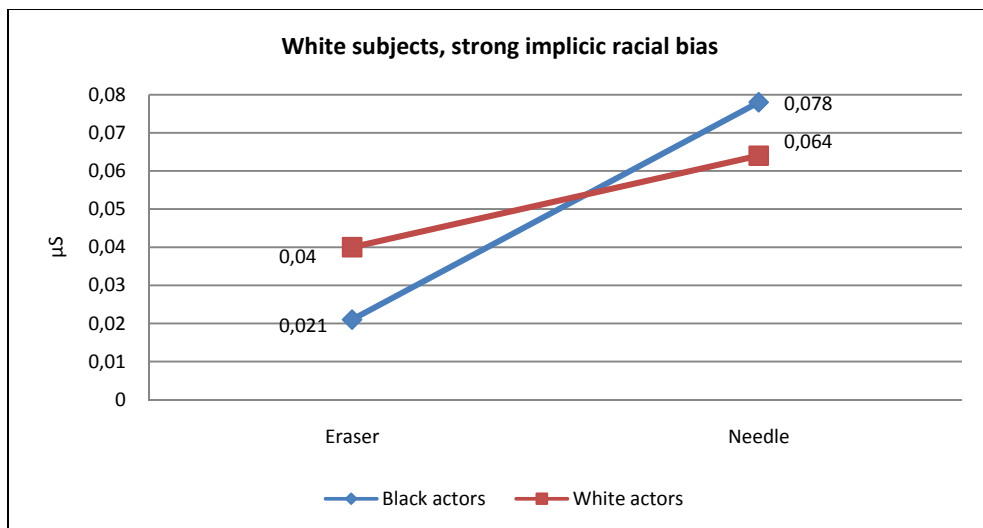


Fig. 24 Data show the empathic reactions of Whites with high IAT scores for White and Black actors

The BEES empathy scale and the 4 subscales of the IRI empathy scale did not show any significant effects in the model (all  $p > 0.05$ ). Estimating the stimulus x actors' ethnicity x actors' gender x subjects' gender model on Black subjects' data revealed that Black subjects did not show any empathy reactions neither for White actors nor for Black actors. The stimulus main effect is not significant  $F(1,15)=0.54$ ,  $p=0.47$  as well as the stimulus x actors' race interaction  $F(1,15)=0.29$ ,  $p=0.59$ . Fig. 25 shows data for the interaction stimulus x actors' race.

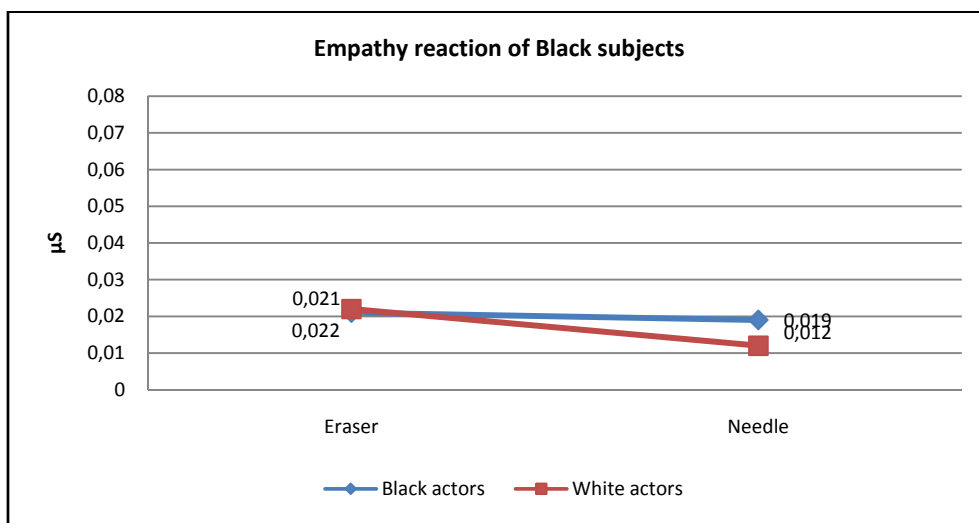


Fig. 25 Data show that Black subjects did not react to the pain of other people. The lack of reaction is independent from the race of the person in pain.

In the black sample the interaction stimulus x actors' race x IAT is not significant  $F(1,12)=0.97$ ,  $p=0.34$ .

Fig. 26 shows that black subjects with low IAT scores (scores below the sample average) did not show a different empathy for black and white actors  $F(3,12)=0.64$ ,  $p=0.6$ .

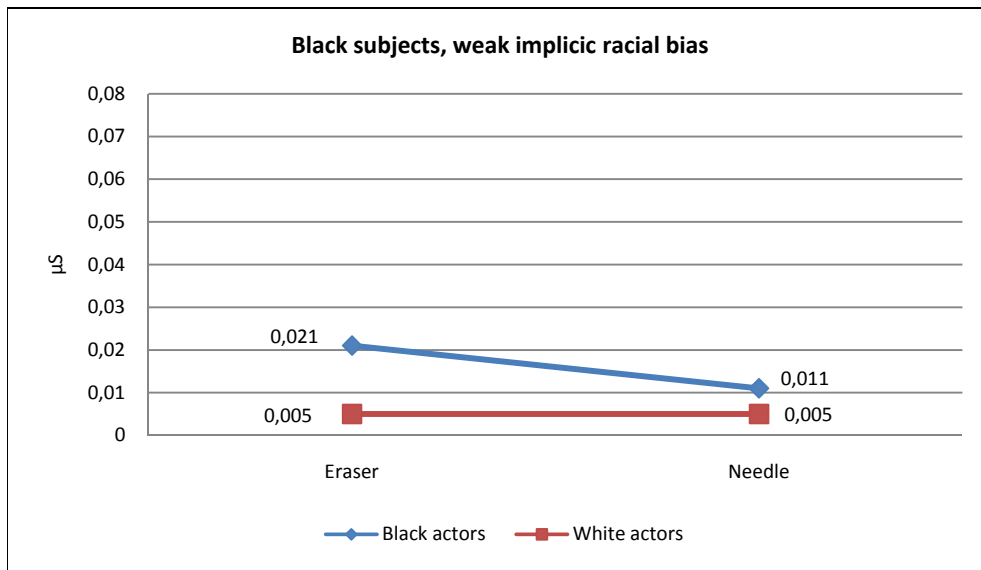


Fig. 26 Data show that black subjects with a low negative implicit bias for Blacks did not show any differences in empathy for Blacks and Whites.

Subjects with high IAT scores (scores above the sample mean) showed the same pattern of data revealing no interaction between stimulus reaction and actors' race  $F(3,12)=1, p=0.42$  (see Fig. 27)

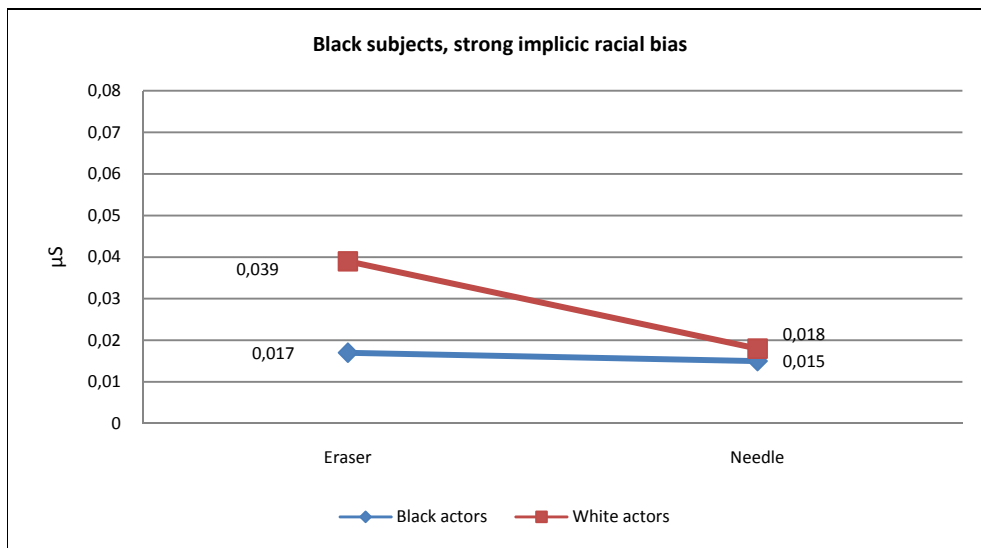


Fig. 27 Data show that Blacks with a strong negative implicit bias for black people had the same reaction for the pain of Whites as for the pain of Blacks.

## **Discussion**

Taking together this data indicate that white people living in the Boston area do not show any differences in the empathy reaction for Black or for White actors. The implicit measure of prejudice against blacks does not seem to be able to shed light on this effect. Although the interaction stimulus x actors' race x IAT is significant data are not consistent with predictions and critically they do not have a structure that leads to alternative meanings. People with a stronger negative bias for Blacks seem to react more for the pain of Blacks than for the pain of Whites. The similar reaction for the pain of Whites as for the pain of Blacks is not consistent with data I collected in Milan with a very similar paradigm. This could be due to the differences in the two different social contexts. Boston is a very multicultural city and Harvard is one of the most multicultural university in the world. It could be possible that white students living in Boston and attending Harvard classes do not see Blacks different from Whites. Slavery has never been allowed in Massachusetts, even at the very beginning of North America modern history. Boston is known as the city where the fight of independence of American colonies started. Equal rights for different peoples were one of the most important claim people were fighting for. Massachusetts is likely to be one of the most liberal State in United States of America and is considered a "blue" State since Democrats have ruled it for years and years. Whites students in Boston area are likely to reflect this culture and they are likely to perceive Black people and White people with very weak cognitive bias.

On the other hand Milan is not yet a multicultural city as Boston is now. In Milan there are less Black people in percentage than in Boston. In north of Italy gained more and more consensus political parties that sometimes make claim against black people and, broadly speaking, they make claims against people



coming from different Counties or different cultures. My data collected in Milan demonstrate that the lack of empathy for Blacks can not merely be due to the different race of black people (Whites subjects did not show lower empathy for Chinese people). Whites subjects in Milan are likely to be part of a different cultural environment than whites subjects in Boston. Social differences between Milan and Boston may explain the different empathy reactions showed by American and Italian subjects.

While Whites showed empathy reactions, Black subjects in Boston did not show any reactions for the pain of other people. As previous research showed, African Americans have lower Skin Conductance levels than Caucasian Americans. In 2006 Brown and colleagues (Brown et al., 2006) found that African Americans were less likely than European Americans to have a skin conductance response to a series of pleasant and unpleasant pictures portraying either ingroup or outgroup members. Other research suggests that such differences are due to skin color differences: regardless of people's race, darker skin has less electrodermal conductivity than lighter skin (Korol & Kane, 1978).

At the same time new videos used in Boston are slightly different from videos used in Milan. In Boston I made 40 new videos very similar to Milan videos as for the shot and timing. But I used a syringe different from the one I used in Milan. We bought a fake syringe, smaller than the Milan one, with a fake needle that can enter the syringe so that it does not hurt the skin of the subjects.

It's possible that subjects perceived this stimulus as less painful than the syringe I used in Milan. Considering the lower basal level of skin conductance in Black people, it's possible that this new stimulus could be enough painful and stressful to trigger physiological reactions in Whites but not in Blacks.

Taking together data gathered in Milan and in Boston indicate that Skin Conductance provides a good index of vicarious reactions for pain of other people.

For the first time I showed that Caucasians people share other people's pain according to the race of the person who is in pain. These data shed light on the importance of social cues on emotions recognize. But most important my data indicate that while we are fully not aware our body perceives social cues and reacts according to them on the base of internalized associations. As discussed above, the implicit social cognition predicts such a behavior and it allows us to ground them in the Reflective system (Strack & Deutsch, 2004). As all of us know, pain perception is one of the most important social emotions and triggers not only helping behavior but also many choices we made every day in our family, during our job, in the street. The question we ask us is "what is her/his race?". We do not ask us: "Who he/she is?". Stereotypes seem to be pervasive. Since 1950s (Allport, 1954) we have known stereotypes do affect our mind, my data clearly show that stereotype affect our brain as well.

# **Racism and empathy for pain: an fMRI experiment**

## **Introduction**

Although pain has been considered an intimate and private feeling, neuroimaging data gathered with fMRI techniques, indicate that when people witness or imagine the pain of another person, they map the others' pain onto their brain using the same network activated during firsthand experience of pain as if they were vicariously experiencing the observed pain (Bufalari et al., 2007; Carr et al., 2003; Cheng et al., 2008; Hutchison, Davis, Lozano, Tasker, & Dostrovsky, 1999; Lamm, Batson et al., 2007; Morrison et al., 2004; Singer et al., 2004; Wicker et al., 2003).

The expertise to understand and to share others' people feeling is fundamental in everyday life and affects most of decisions, behaviors and emotions that shape our single day. And crucially for our social interactions, it has been demonstrated that empathic reactivity to others' feelings and pain is deeply affected by social cues and individual differences. We now know that functional activity related to empathy reactions to others' feelings is affected by similarity between the witness and the person in pain (Krebs, 1975; Lamm, Meltzoff, & Decety, 2010; Preston & de Waal, 2001), by previous experience in the same situation (Cheng et al., 2007), by observer's personality (Avenanti, Minio-Paluello, Bufalari, & Aglioti, 2009), by familiarity for the target (Cialdini et al., 1997), by gender (Baron-Cohen & Wheelwright, 2004; Eisenberg & Carlo, 1995), and age (Eisenberg & Morris, 2001) and, most relevant for the present work, initial findings seem to suggest that empathy for others pain can be modulated by ingroup / outgroup social categorization based on race differences between the

target and the onlooker (Avenanti et al., 2010; Cosmides et al., 2003; Xu et al., 2009).

Despite the growing body of research on this topic, little is known on the relation between functional brain activity related to empathy for pain and implicit bias towards outgroups members. With the present work I seek to determine whether and how functional brain activities in response to the pain of other people are affected by implicit, non-conscious negative prejudice for members of a different race.

More specifically, I aim at understanding whether the implicit negative bias for the outgroup affects the activation pattern among brain regions that vicariously respond to other people's pain. Since sharing other people's feelings affects our everyday behaviors, it's of great importance to understand whether non-conscious negative bias affects how low-level perception brain regions activate brain regions deputed to elaboration of high-level emotions and thoughts. It may be that internalized associations for specific social groups affect how our brain works and how feelings of other people are elaborated. My previous experiments showed subjects had a different Skin Conductance in response to pain of people of different races. With this experiment I want to understand why they do that and where is the origin of that difference in physiological activity.

I chose functional Magnetic Resonance Imaging Technique to test my hypothesis because this neuroimaging technique allows us to precisely measure how the brain works in specific experimental conditions.

## **Stimuli**

Forty videos have been made. Each video started with a frame depicting a face of a female or a male actor holding a neutral expression. Subsequently the

camera zoomed in on the actor's hand which was touched by the experimenter alternatively by an eraser (harmless stimulus) or by a needle (painful stimulus). Twenty actors have been recruited: five black males, five black females, five white males and five white females. Each video lasts for twelve seconds. During the first 8 seconds the scene is playing and at the end of the 8<sup>th</sup> second the stimulus (alternatively the needle or the eraser) is touching the hand of the actor. On that precise frame the video stops playing and a still image with the very last frame keeps playing for the next 4 seconds. Then for each stimulus the subjects are presented with an 8 seconds video and a 4 seconds still image.

## **Method**

During the fMRI sessions, stimuli were projected from a PC located outside the MR room and connected via optical fibers to dedicated goggles (Visuastim XGA, Resonance Technology, [www.mrvideo.com](http://www.mrvideo.com)), using Presentation 11.1 software.

The 40 video clips were presented in a random order. After each movie a question mark appeared on the centre of the screen of the computer and lasted for 3 seconds. During this period of time, participants were asked to judge how painful was the situation represented into the video clip using a Likert scale from 0 (not painful at all) to 3 (highly painful). To do this they were instructed to press a button on a five-buttons keyboard where each key corresponded to a value of the "pain-scale".

To avoid artifacts due to mouth and head movements participants were also instructed not to move their head and not to talk.

## **Behavioural measures**

Before the scan session, participants were placed in a safe and quiet room and were asked to complete on a desktop pc the Mehrabian trait empathy scale (Mehrabian & Epstein, 1972). After completing the explicit measures session, subjects were asked to complete a race Implicit Association Test to assess the implicit race bias for Black people with respect to White people (Greenwald, McGhee, & Schwartz, 1998). The computer based session took no more than 20 minutes.

## **fMRI data acquisition**

MRI scans were performed on a 1.5 T Marconi-Philips Infinion Scanner, using an Echo Planar Imaging (EPI) gradient echo sequence (Flip angle = 90°; TE = 60msec; TR = 3050 msec; FOV = 240x240; matrix = 64 x 64). The selected volume consisted of 26 contiguous transverse images (thickness = 5 mm; gap = 0 mm), acquired every 3.05 seconds. 214 scans were collected for each participant.

## **fMRI analyses**

The fMRI analyses were performed using the statistic software SPM2 (Wellcome Department of Cognitive Neurology).

A standard pre-processing, including realignment, normalization and smoothing with a Gaussian filter of 10x10x10 mm, was performed.

After this, the four experimental conditions for each subject were modeled in an event-related design. In particular the BOLD signal was convolved with a standard HRF (hemodynamic response function) (Friston et al., 1995) and adjusted for global differences using a proportional scaling for all voxels. Moreover, using a high pass filtering technique, the possible confounding

contributions to the fMRI signal coming from cardiac and respiratory cycles were removed.

For each participant a fixed-effect analysis was performed in order to generate individual contrast images (con-images) containing the voxel-by-voxel information about the effect size of neutral and painful stimulation in Caucasian and African actors separately.

These images were used to design a 2x2 second-level ANOVA (Friston, Stephan, Lund, Morcom, & Kiebel, 2005) which allowed us to isolate the neural substrate associated with observing a painful stimulation (empathy for pain condition). In particular, the direct comparison between painful and neutral stimulation was calculated using an inclusive masking technique to restrict the number of comparisons only to the voxels that were activated during the task (i.e. the main effect of task irrespective from the experimental condition). The threshold for the comparison was set at  $p < .001$ , while the threshold for the inclusive masking procedure was set at  $p < .05$ .

Once isolated the brain regions activated during the empathy for pain condition, the results reported in the review by Peyron and colleagues (Peyron et al., 2000) were used to construct an anatomical mask including the brain regions typically subserving pain perception, i.e. the anterior cingulated cortices, the postcentral gyri, the Rolandic opercular gyri, the insulae and the thalami. We refer to this pool of brain regions as “pain matrix”. The anatomical Region of Interest (ROI) corresponding to the areas described in the paper by Peyron and colleagues were isolated using the AAL template (Anatomical Automatic Labeling template) (N. Tzourio-Mazoyer et al., 2002) available with the software MRICro (Rorden C & Brett M, 2000) available at [www.sph.sc.edu/comd/rorden/mricro.html](http://www.sph.sc.edu/comd/rorden/mricro.html)

Once isolated, each anatomical region was transformed as ROI and by means of the “union” function we created a single ROI containing all the voxels belonging to the *pain matrix*. The *pain matrix* ROI was then transposed to the custom-fMRI template created in our laboratory in order to have an MNI stereotactical space of the same dimension of our fMRI data. This operation allowed us to save the transposed *pain matrix* ROI as a new ANALYZE file which was then used as volume of interest for the Small Volume Correction analysis.

The pain matrix was used as volume of interest in a small-volume correction (SVC) analysis to assess whether the brain regions associated with pain perception may have a role also when experiencing another’s pain. (see Fig. 28 later in the text)

Finally, to evaluate whether the ethnic group and a measure of implicit attitude for Blacks and Whites (Greenwald et al., 1998) may modulate the brain response to empathy for pain in regions belonging to the pain matrix, the BOLD signal extracted from the activation peaks surviving the SVC analysis were entered in a 2x2 within subjects ANCOVA where the standardized IAT score for each participant was introduced as covariate variable. This analysis was performed using the software package SAS.

## **Subjects**

For this study, 16 participants, 8 males and 8 females (mean age=25,3 years, SD=4,81) were recruited among undergraduate students of Milan Bicocca University. None of the participants had any history of neurological disorders or learning disabilities. All the participants were right-handed.



## Results

### Behavioral results

The sample showed a significant negative bias for Blacks. The IAT mean= 0.53 (SD=0,41) is significantly greater than 0,  $T(15)=5.18$ ,  $p<0.0001$ . This expected result indicates that subjects showed on average a significant negative bias for Black people over White people.

In the sample the BEES scales showed a mean of 167 points (SD=24).

The explicit rating of the pain perceived during the videos showed that subjects did not perceive any differences in pain intensity between stimuli with Whites and stimuli with Blacks (Tab. 4). The paired T-test between painful condition on Blacks and painful condition on Whites confirmed these data  $T(14)=1.07$ ,  $p=0.3$ .

	Mean	N	Standard deviation	Standard Error
Needle Blacks	1,530	15	,6313	,1630
Needle Whites	1,565	15	,6116	,1579
Eraser Blacks	,007	15	,0287	,0074
Eraser Whites	,221	15	,1065	,0275

Tab. 4 Means of average rating during the scanning procedure for videos containing painful stimuli on Whites and Blacks and harmless stimuli on Whites and Blacks.

### fMRI results

The neural network activated by our experimental task, irrespectively of the experimental conditions (i.e. the main effect of task), included the prefrontal cortices, the primary and secondary somatosensory cortices bilaterally, the right insula, the neocortical regions of the temporal lobe bilaterally, and medial temporal structures such as the right hippocampus and the two amygdalae. (see Tab. 5)

Brain regions	Left hemisphere				Right hemisphere			
	<i>MNI Coordinates</i>							
	<i>x</i>	<i>Y</i>	<i>z</i>	<i>Z score</i>	<i>x</i>	<i>y</i>	<i>z</i>	<i>Z score</i>
	<b>Main effect of task</b>							
Mid. frontal gyrus					44	48	2	5.2
					42	46	10	5.1
Mid. frontal orb. gyrus	-40	54	-2	5.0	32	52	-12	5.5
					36	50	-14	5.5
Inf. frontal orb. gyrus	-40	52	-14	5.1				
	-42	52	-10	5.0				
Rolandic opercular gyrus	-64	-10	14	5.0	60	-12	14	5.2
Insula					42	4	-2	5.6
					38	12	-10	4.8
Amygdala	-24	0	-18	5.5	28	2	-14	5.5
Hippocampus					22	-4	-20	4.9
Postcentral gyrus	-48	-20	48	7.3				
	-44	-40	58	5.5				
Inf. parietal gyrus	-42	-46	58	5.7	52	-36	58	6.8
	-48	-38	56	5.3	48	-50	56	6.7
Angular gyrus					40	-66	46	5.1
					42	-56	50	4.9
Sup. temporal pole	-54	12	-10	6.8	50	16	-14	6.7
					56	8	0	5.5
Sup. temporal gyrus	-46	4	-6	7.4	66	-18	16	5.8
	-64	-20	12	5.7				
Mid. temporal gyrus	-64	-48	-10	5.3				
Inf. temporal gyrus	-60	-54	-16	6.3	60	-42	-12	5.5
Cuneus					2	-84	26	6.0
Inf. occipital gyrus					36	-94	-8	6.7
					26	-100	-8	5.2
Cerebellum	-44	-64	-36	7.3	40	-52	-34	5.7
	-30	-36	-28	5.1	38	-48	-34	5.6

Tab. 5 Main effect of task. Only FWE-corrected results are reported ( $p < .05$  FWE-corrected)

Among these regions, the left inferior frontal gyrus (par orbitalis), the left cerebellum, the right insula, the right supramarginal gyrus, the right inferior parietal gyrus, the right inferior occipital gyrus and the postcentral gyri bilaterally were more activated during the empathy for pain condition when compared with the neutral condition (see Tab. 6).

Brain regions	Left hemisphere				Right hemisphere			
	<i>MNI Coordinates</i>							
	<i>x</i>	<i>Y</i>	<i>Z</i>	<i>Z score</i>	<i>x</i>	<i>y</i>	<i>z</i>	<i>Z score</i>
Inf. frontal orb. gyrus	-44	44	-12	3.7				
Insula *					36	20	-10	3.6*
Postcentral gyrus *	-54	-20	26	4.5*	66	-16	34	4.4*
					48	-30	48	3.2*
Supramarginal gyrus					68	-16	32	4.6
Inf. parietal gyrus					40	-42	50	4.6
Inf. occipital gyrus					34	-94	-4	3.6
Cerebellum	-30	-72	-34	4.0				

Tab. 6 Brain regions showing a significant effect of painful stimulation versus neutral stimulation. \* = Brain regions belonging to the pain matrix. These regions had been identified by means of a SVC approach.

Part of these activations fell within the so called pain matrix; in particular a significant overlap was found in the postcentral gyri bilaterally and in the right insula (see Fig. 28).

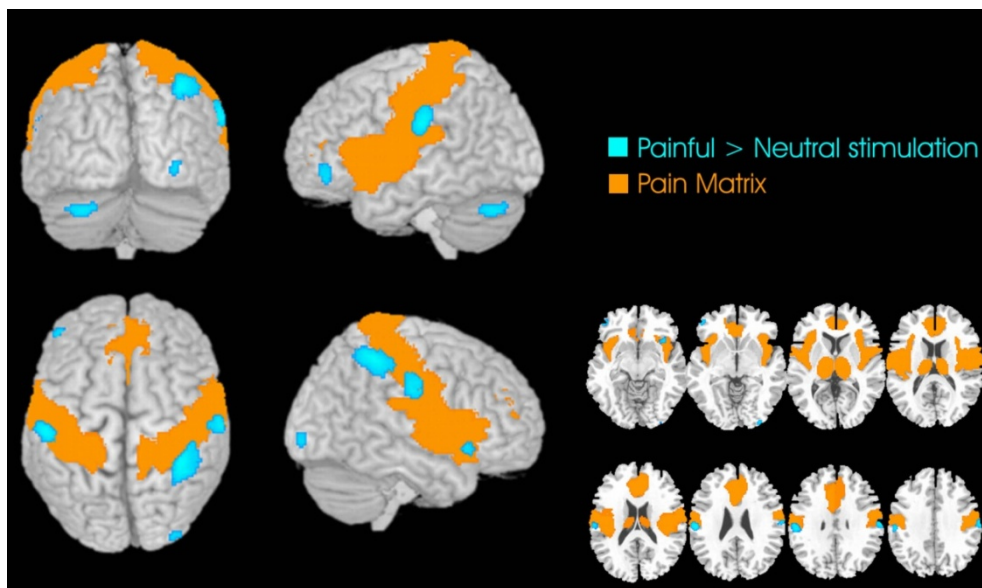


Fig. 28 Figure shows areas involved in the pain matrix (Peyron et al., 2000) (orange areas) and areas that on the base of our data resulted significantly more activated in painful condition than in neutral condition (blue areas).

To test our hypothesis the BOLD signal extracted from the activation peaks surviving the SVC with the pain matrix was further explored with a 2 (race of the target) x 2 (type of stimuli) repeated measures ANCOVA model with the IAT

subjects' scores as a covariate. Since we have clear hypothesis on the direction of the effects, we would not interpret effects against our hypothesis as significant. Hence we interpreted all the effects in the model as one tail tests.

The ANOVA model confirmed that painful stimuli triggered a significantly stronger activity in right insula than harmless stimuli  $F(1,14)=35,2$ ,  $p<0.0001$ . At the same time data revealed that this effect is not affected by the race of the actors  $F(1,14)=0.31$ ,  $p=0.58$ . The main effect of actors' race is not significant  $F(1,14)=0.19$ ,  $p=0.66$  meaning that the different race of the actors did not trigger any differences in insula activity. As expected, the IAT main effect is not significant  $F(1,14)=0.69$ ,  $p=0.42$  meaning that IAT scores did not predict activations in insula. At the same time the interaction IAT x targets' race is significant  $F(1,14)=6.34$ ,  $p=0.01$  ( $p=0.02$  two tails). Fig. 29 shows that subjects with a strong negative bias for Blacks, i.e. high IAT scores, show a insula activity higher for Whites than for Blacks.

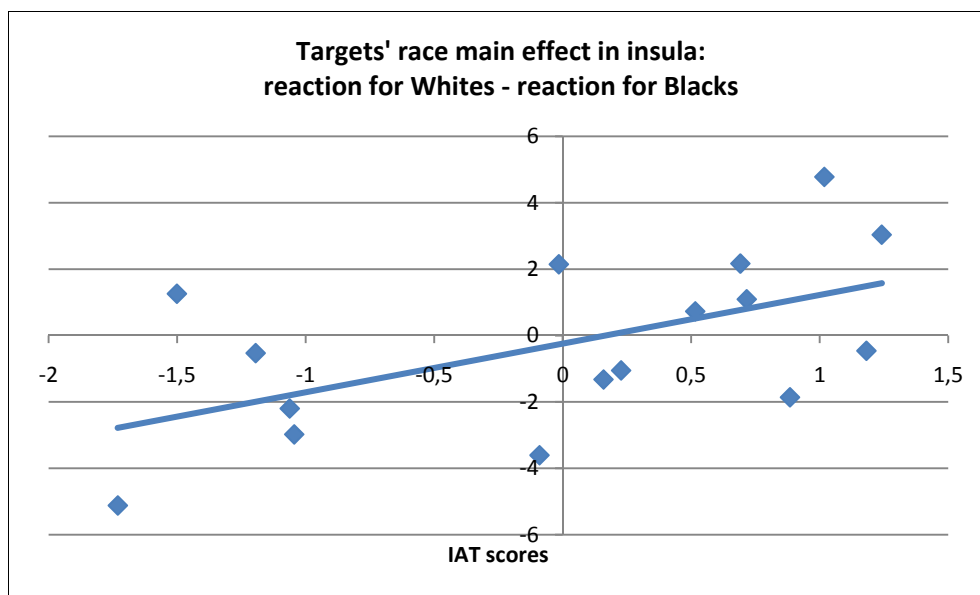


Fig. 29 Data show the significant linear effect of IAT scores on actors' race effect in Insula. The linear trend indicates that the higher the IAT scores, the higher the difference between the reaction for Whites compared to reaction for Blacks.

Critically the three way interaction stimulus x targets' race x IAT is significant  $F(1,14)=4.07$ ,  $p=0.03$  (0.06 two tail). This means that the implicit negative bias for Blacks affected the difference between the insula reaction for the pain of Blacks and reaction for the pain of Whites. Specifically we estimated a simple slope analysis (Aiken & West, 1991) to estimate the stimulus X targets' race interaction for subjects 1SD above the sample mean, for subjects 1SD below the sample mean and for the average subject (see also previous chapter for a similar analysis). Fig. 30 shows that subjects with a strong negative implicit bias for Blacks (IATHigh) showed a stronger reaction in insula for pain of Whites than for pain of Blacks. Subjects with a low race bias (IATLow) reacted more for the pain of Blacks than for the pain of Whites. The average subject (IATMedium) reacted slightly more for the pain of Blacks than for the pain of Whites.

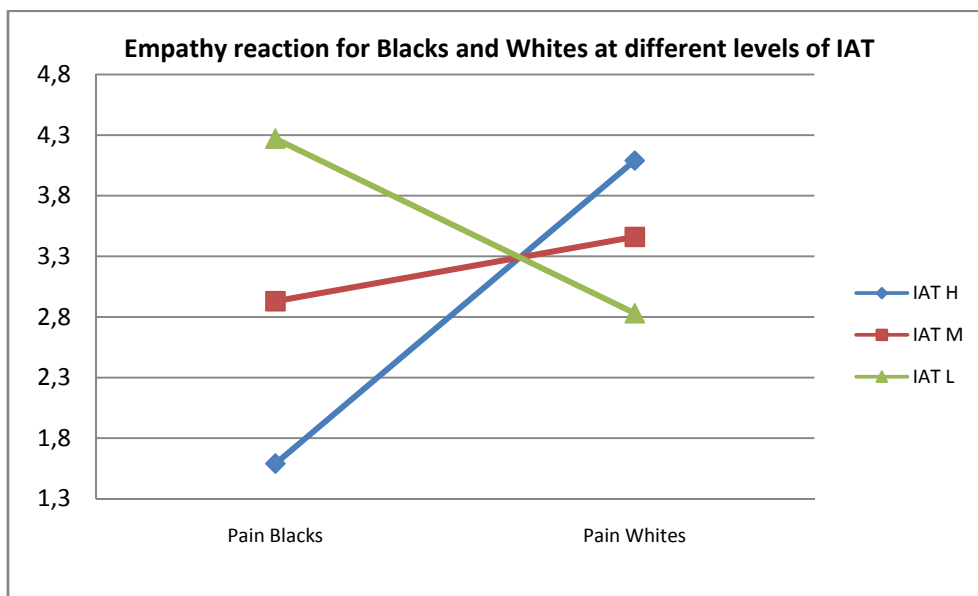


Fig. 30 Simple slope analysis for the three way interaction Pain x race x IAT. Data show that subjects with a strong negative bias for Blacks with respect of Whites (IAT scores 1 SD above the sample mean, IATH) reacted more for the pain of Whites than for the pain of Blacks. The sample on average (IATM) reacted slightly more for the pain of the race ingroup. Subjects with a weak negative bias for Blacks (1 SD below the sample mean, IATL) reacted more for the pain of Blacks.

The BEES empathy scale did not show any significant effects in the stimulus X actors' race X BEES ANCOVA model. All  $p > 0.05$ .

To test our hypothesis on whether the implicit bias for the outgroup affects the pattern of vicarious activation among regions of the pain matrix, we estimated a series of path analysis models.

We included in the models the IAT scores, activation in left postcentral gyrus (left somatosensory 1 area), right postcentral gyrus (right somatosensory 1 area, two different peaks, see Tab. 6), interaction between activations in these regions and the IAT scores and we estimated whether using these variables we can predict the different activation in the insula for White with respect of Blacks.

We estimated three different models: a model that estimates the effects of the empathic reaction in the left somatosensory I area (the difference between the activation in painful and harmless condition) and the empathic activation in right somatosensory I area (two different peaks) on the empathic reaction in the insula area.

We estimated such a model using data of activations for stimuli with black actors and a different model with data of activations for stimuli containing white actors.

The third model contains data of functional empathic activations during the task, with no difference between stimuli containing black or white actors.

These models show how activations in low level perception areas (left and right S1) affect activation in an high level cognitive area (insula).

Fig. 31 shows that considering reactions for white actors, data indicate that empathic reaction in primary left somatosensory area predicts empathic activation in insula. Empathic activations in primary somatosensory areas do not predict empathic activations in insula.

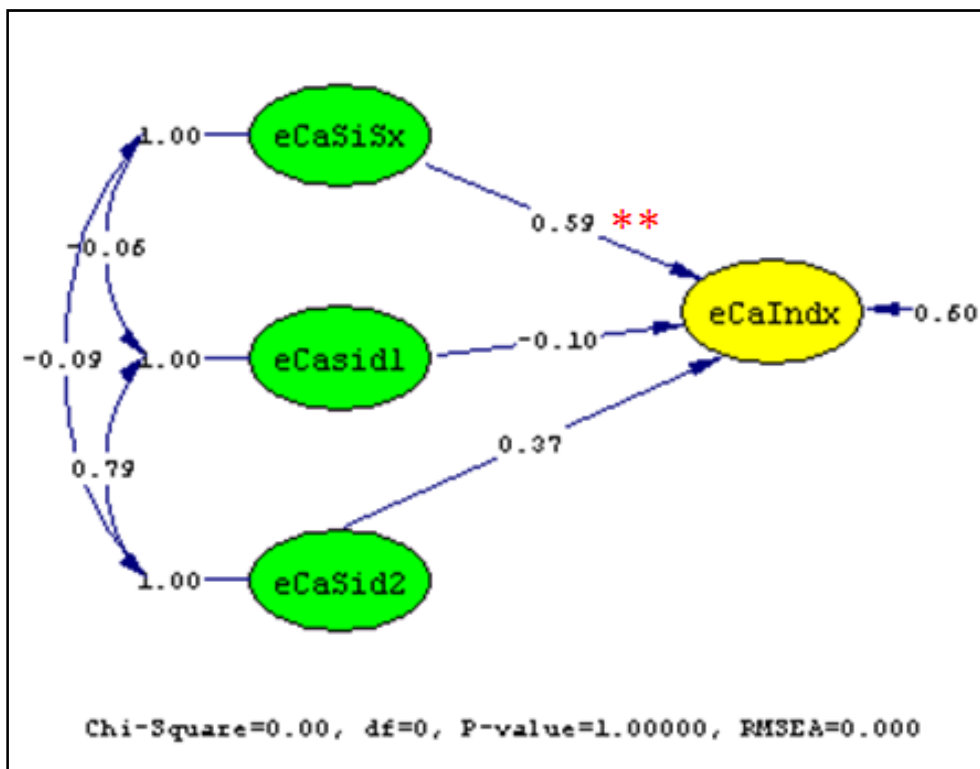


Fig. 31 Functional activations for white actors. Values are standardized effect among variables. Straight arrows indicate direct linear causal effects. Round arrows indicate correlations. eCaSiSx=Empathic reaction (painful – harmless condition) in left S1. eCaSid1= Empathic reaction (painful – harmless condition) in right S1, first peak. eCaSid2= Empathic reaction (painful – harmless condition) in right S1, second peak. eCaIndx= Empathic reaction (painful – harmless condition) in right insula. \*= $p < 0.05$  \*\*= $p < 0.01$

Considering that actors in videos were always touched on the right hand, data indicate that when subjects were watching at videos with whites actors, the empathic activation in controlateral primary somatosensory area significantly affected the activation in the insula area. Empathic activation in ipsilateral primary somatosensory area did not predict empathic activations in insula.

I estimated the same structural equation model using data of activations for black actors. Fig. 32 shows that when subjects were watching at videos with black actors, empathic activation in left primary somatosensory area does not predict empathic activation in insula. At the same time empathic activation in one of the two peaks in the right primary somatosensory area significantly affected the

empathic reaction in insula with a standardized effect of 0.49. Correlations among the S1 area were not significant meaning that S1 area did not activated with the same pattern during perception of the pain of Blacks.

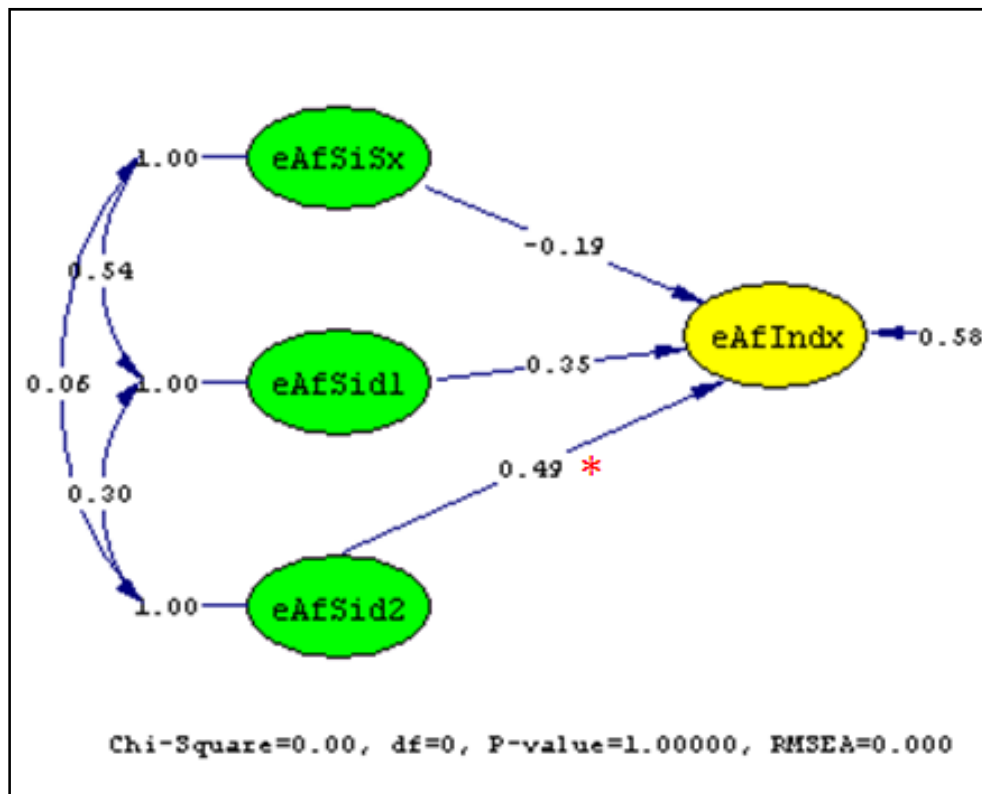


Fig. 32 Functional activations for black actors. Values are standardized effect among variables. Straight arrows indicate direct linear causal effects. Round arrows indicate correlations. eAfSiSx=Empathic reaction (painful – harmless condition) in left S1. eAfSid1= Empathic reaction (painful – harmless condition) in right S1, first peak. eAfSid2= Empathic reaction (painful – harmless condition) in right S1, second peak. eAfIndx= Empathic reaction (painful – harmless condition) in right insula. \*= $p < 0.05$  \*\*= $p < 0.01$

Taking together data of activations for stimuli containing black actors show that empathic activation in the contralateral primary somatosensory area does not trigger the empathic activation in the insula. At the same time activations in one of the two peaks in the ipsilateral primary somatosensory area does predict empathic activations in insula.

To further explore our data I estimated a path analysis model containing average brain activations for black and white actors. I included in the model the



IAT subjects' scores, the empathic reaction in left primary somatosensory area and right primary somatosensory area (2 different peaks), interactions between IAT scores and empathic activation in left and right SI areas. The model estimates which variables predict the difference between insula empathic activation for whites and for blacks (hereinafter "I3": Insula Interaction Index): in this variable positive values indicate higher empathic reactions for whites than for blacks. Fig. 33 shows that the empathic reaction in left SI significantly trigger the I3 with a positive correlation. The IAT scores significantly predict the I3 and data indicate that the higher the IAT scores, the higher the I3. Furthermore, since the interaction between the empathic reaction in right SI (peak II) and the IAT does affect I3 values, taking together data indicate that empathic reaction in right S1 area (second peak) affects the I3 depending on the IAT scores. Specifically high IAT scores (strong negative bias for blacks) make stronger the effect of right S1 on I3. All the correlation among the independent variables are not significant.

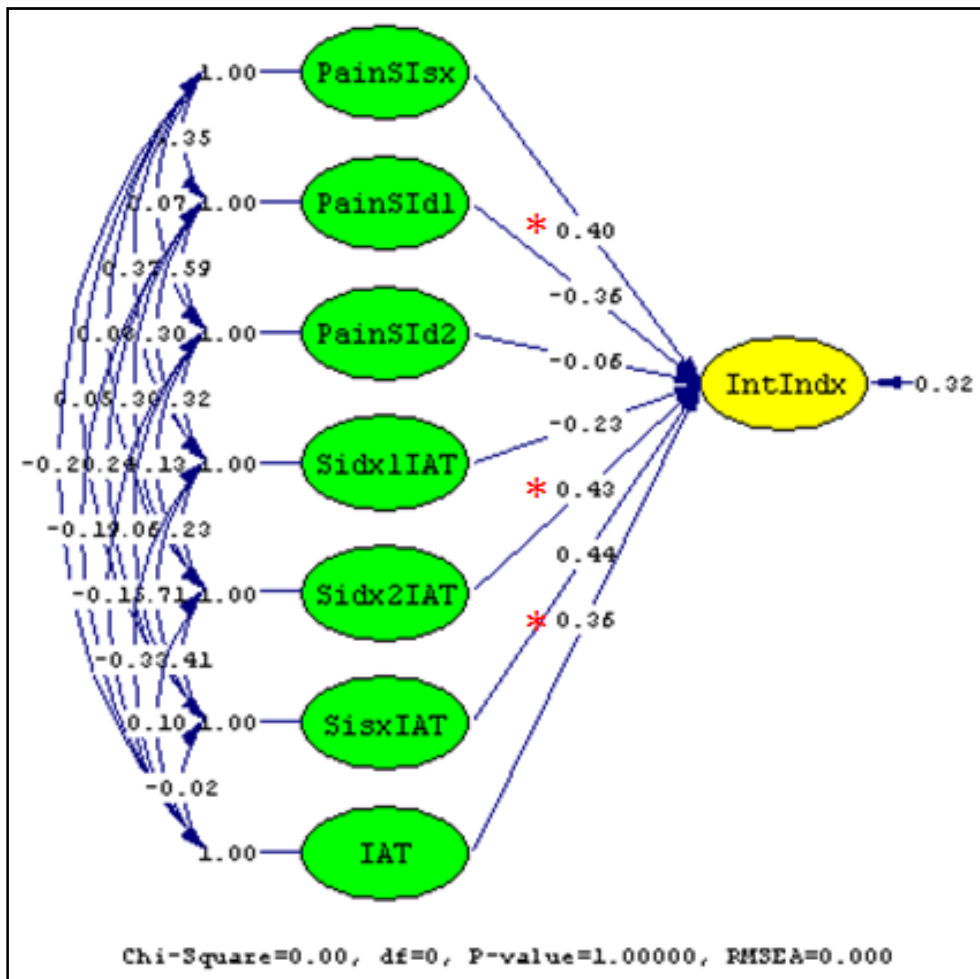


Fig. 33 Functional activation during the pain perception task. Values are standardized effect among variables. Straight arrows indicate direct linear causal effects. Round arrows indicate correlations. PainSiSx=Empathic reaction (painful – harmless condition) in left S1. PainSid1= Empathic reaction in right S1, first peak. PainSid2= Empathic reaction in right S1, second peak. Sidx1IAT=interaction effect between empathic reaction in right S1 first peak and IAT scores. Sidx2IAT=interaction effect between empathic reaction in right S1 second peak and IAT scores. SisxIAT=interaction effect between empathic reaction in left S1 and IAT scores. IAT=Implicit race association test scores. High values indicate strong negative bias for Blacks. \*= $p < 0.05$  \*\*= $p < 0.01$

## Discussion

My data indicate that painful stimuli triggered significant activations in area previously known to be involved in vicarious pain perception. Bilateral primary somatosensory areas and right insula were significantly more activated for painful stimuli than for harmless stimuli irrespectively for the race of the person in pain. But critically my data show that vicarious activations in insula for other people

pain are moderated by the implicit negative bias for black people. Analysis I described reveal that an average subject reacts both for the pain of Blacks and for the pain of Whites and insula does not distinguishes between the two races, but IAT scores reveal a meaningful trend in the sample. Subjects with a strong negative bias for Blacks show an insular stronger activity for the pain of whites than for the pain of Blacks. These data are consistent with previous studies that showed that brain activity is moderated by relevant social cues of the person in pain (Singer et al., 2006) or by implicit perception of people similar or dissimilar to the self (Mitchell, Macrae, & Banaji, 2006). Relevant studies showed that implicit negative bias for blacks affects sensorimotor empathic brain responses for black and white individuals (Avenanti et al., 2010), but my data, for the first time, seem to indicate that our brain when is perceiving pain of another person reacts accordingly to internalized, non conscious associations to the person who is in pain. This associations trigger the insula to a weaker activity for pain of people associated with negative internalized social cues.

Path analysis models showed that when subjects perceived a white person in pain, i.e. a ingroup member, controlateral S1 area perceived the pain and triggered insula empathic reaction. These data are consistent with many studies on empathy for pain (Derbyshire, 2000) and specifically on S1 role in pain perception (Bushnell et al., 1999; Keysers, Kaas, & Gazzola, 2010; Talbot JD et al., 1991). At the same time when subjects were watching a black person in pain, i.e. an outgroup member, ipsilateral S1 perceived the pain and triggered the insula empathic reaction. Furthermore the right S1 area seems to interact with interiorized negative bias for blacks in determining the lack of empathic reaction for Blacks in insula.

Taking together our data seem to outline a different role of left and right S1 in triggering insula area: the contralateral S1 triggers insula when subjects put themselves in other person's shoes and mirror his/her pain. Ipsilateral S1 triggers insula when subjects do not mirror other people pain but can perceive the pain of other person.

Our data indicate that empathy for pain is not just affected by the race of the person in pain (Xu et al., 2009), but perception of pain relies on different anatomic paths according to the race of the person in pain. This could lead to encode ingroup members' pain on the base of a self-other overlap. On the other hand outgroups' members' pain could be encoded by more perception and sensory discrimination.

## **Religion and empathy for pain: a pilot study**

In modern multicultural societies, religion, beside race, is one of the most important element that creates social boundaries, that differentiates me and you, us and them. Religion refers to deep and ancient values that altogether shape what we are and what they are.

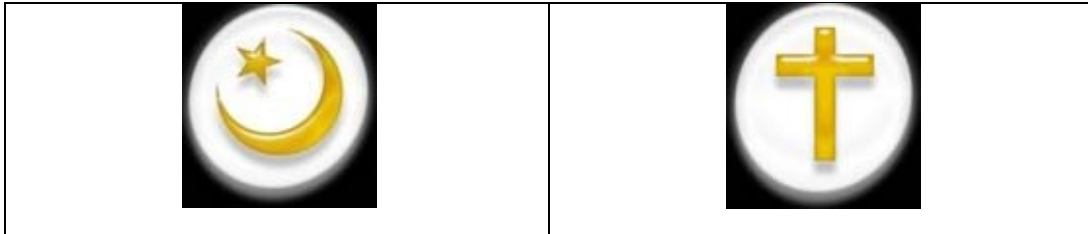
Most important, this is likely to be true in all of modern societies, even in secularized ones.

For sure, religion is a key social value in Italy and it's very likely to shape what we are and what we say we are either if we are believers or we are not.

That's why I choose religion as the second social factor – beside race – I wanted to investigate in order to understand its relevance in shaping how we share feelings of other people.

In the pilot experiment I conducted, I used the SCR to measure empathic reaction for the pain of other people. Forty-three subjects (29 females and 14 males, age: mean=24.7, SD=6.14) have been recruited among undergraduate students of Milano Bicocca University. During the experiment, subjects have been presented with a series of video clip very similar to those used in the previous experiments. In this experiment symbols of Christian religion (a cross) and of Islamic religion (an half moon with a star) have been presented paired with the videos. Before the experiment began, subjects read a set of instructions where they were told they were about to watch a series of videos with one person per video. A symbol would have indicated the religion the person is a believer of. Specifically, on a computer monitor using E-prime software I presented in the center of the screen the videos in a random order. While the videos were playing

in the four corners of the screen were placed the symbols of one religion for each video (Tab. 7).



Tab. 7 Symbols used to indicate Islamic religion (on the left) and Christian religion (on the right).

Each video started with a frame depicting a face of a female or a male actor holding a neutral expression. Subsequently the camera zoomed in on the actor's hand which was touched by the experimenter alternatively by an eraser (harmless stimulus) or by a needle (painful stimulus). A total of 16 videos have been made. All of the actors in the videos looked like Italian Caucasian people and no symbols of any religions were visible in the scenes of the videos. Four male actors and four female actors have been recruited. Each actor have been touched on the hand with both of the stimuli (one for each video). Half of the actors have been paired with the symbol of Christian religion and half of the actors with the symbol of Muslim religion.

Before watching the videos, subjects completed a religion Implicit Association Test (Greenwald et al., 1998) that measured the implicit attitude for Muslims compared to attitude for Christians. Subjects have been recruited among undergraduate students of Milan Bicocca University. Stimuli have been presented in a random order with a 2 (religions, Christian and Islamic) X 2 (actors gender, male and female) X 2 (stimuli, needle and eraser) X 2 items (1st item and 2nd item) within subjects design.

While subjects were watching the video clips, SCR has been collected.

Data show a significant main effect of stimulus type on the psychophysiological subjects' reaction. The painful condition triggered a significantly higher SCR than painful condition  $F(1,42)=17.57, p<0.0001$ .

Religion didn't moderate the stimulus main effect indicating that subjects showed the same physiological reaction either they were watching a Christian actor in pain or a Muslim actor in pain. Interaction stimulus X religion is not significant  $F(1,42)=0.6, p=0.44$  (Fig. 34).

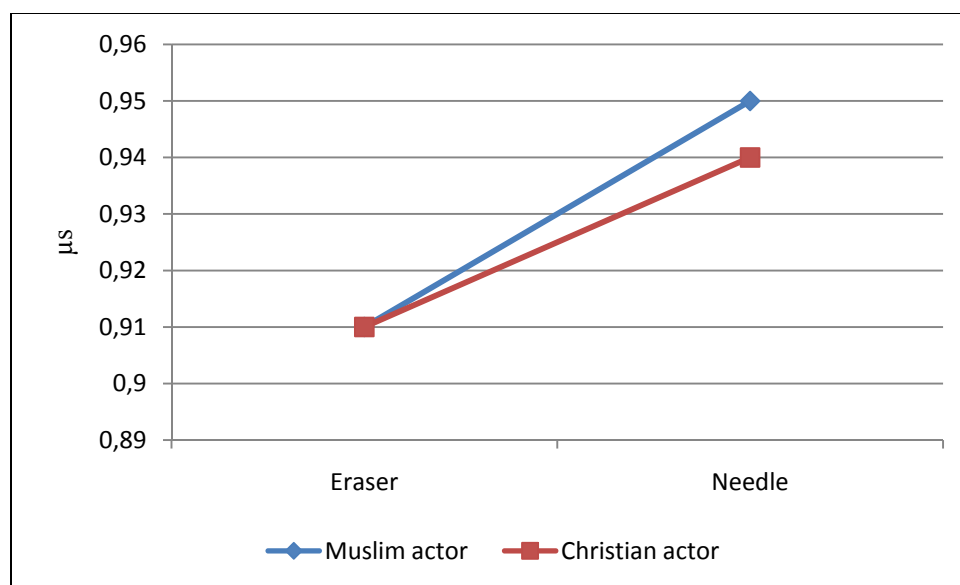


Fig. 34 Data showed that subjects reacted to the pain of Christian actor with the same intensity as the reaction they showed for Muslim actors' pain.

Data showed that the empathic reaction for Christian actors is significant  $F(1,42)=12.31, p<0.001$ . And in a similar vein the empathic reactions subjects showed for Muslim actors is significant  $F(1,42)=5.84, p=0.02$ . These pattern of reaction clearly shows that religion did not moderated the empathic subjects' reaction and indeed subjects reacted equally for the pain of Christian and Muslim actors.

The IAT sample scores (mean=0.56, SD=0.29) are significantly greater than zero  $T(42)=12.77, p<0.001$ , meaning that the sample showed a mean significant

implicit preference for Christian religion over Islamic religion. By estimating the religions X actors gender X stimuli type X times X IAT ANCOVA model data showed that the implicit preferences for Christians over Muslims did not moderate the empathic reactions subjects showed. The IAT main effect is not significant  $F(1,41)=0.35$ ,  $p=0.55$  indicating that, as expected the IAT scores did not predict the SC values of the subjects. At the same time the IAT X stimulus X religion interaction is not significant  $F(1,41)=0.11$ ,  $p=0,74$  indicating that IAT scores did not moderate the interaction stimulus X religion. Fig. 35 and Fig. 36 show data of the stimulus X religion X IAT interaction. Subjects with different implicit bias for Muslims did not show a different empathy for Muslims and for Christians.

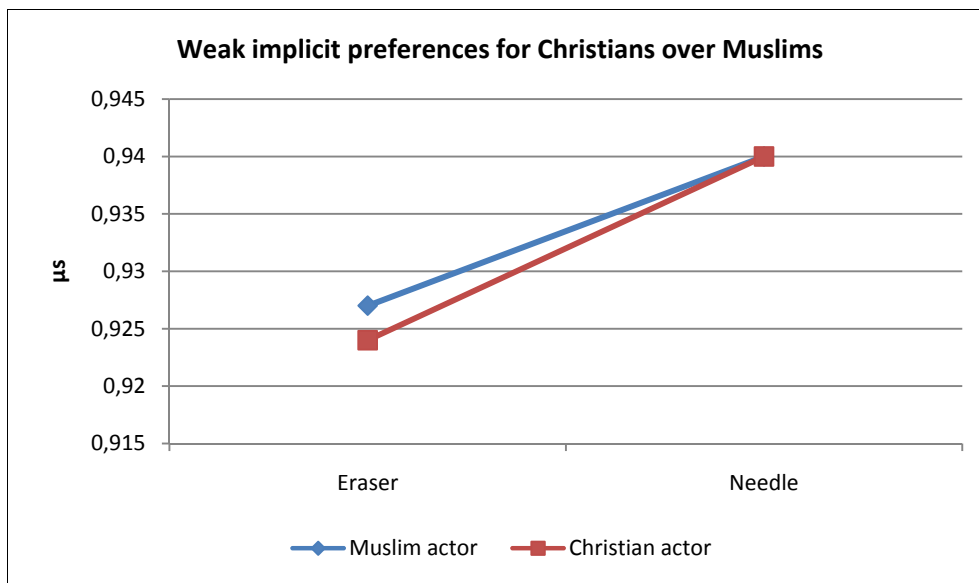


Fig. 35 The figure shows data of the empathic reactions for subjects with a weak implicit preferences for Christians over Muslims.



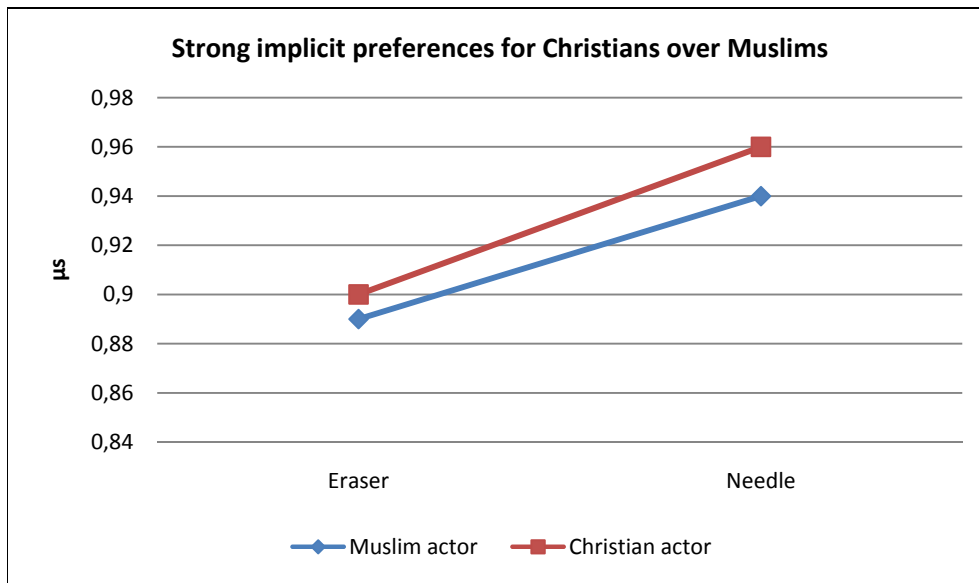


Fig. 36 The figure shows data of the empathic reactions for subjects with a strong implicit preferences for Christians over Muslims.

Data of the pilot experiment showed that subjects did not respond differently to the pain of Christians over the pain of Muslims. And even individual differences in implicit bias towards Muslims did not show any effects on empathic reactions.

Religion of the person in pain seems to be different from race in determining empathic reaction of persons witnessing the pain of other people. Race seems to have a deeper relationship with our body reactions, it seems to play a stronger role in determining automatic reactions. Other researches pointed out that only some social cues are salient in affecting body automatic reactions. In the “general discussion” chapter I will review this kind of researches. Please refer to that chapter for a broader discussion.

## General discussion

The experiments I conducted aimed at researching the relationship between social factors such as race and religion and the brain functioning.

Data I presented showed clearly that our body and our brain are affected by what we think of others and how we perceive them.

Race is a strong moderator of our empathic responses to other people pain, even if we are not aware of doing that. Race creates boundaries that shape the world we interact with, race shapes what we think people are.

But most important, when this mechanism occurs, we are not likely to be explicitly aware of it.

On the other hand, religion does not seem to be able to create such boundaries between me and you, us and them.

But why race is a strong and effective moderator of emotional sharing and religion is not? Answers to this question could shed light on important topics in social psychology and social neuroscience. For example, Batson and colleagues (D. Batson et al., 1997) found that university group membership (i.e., shared or unshared) had no impact on empathetic induction. Conversely, Johnson and colleagues (Johnson et al., 2002) showed that the defendant race clearly affected the empathetic induction on a subsequent juror decision-making task. One reason for the divergent findings might involve the differential nature of the groups studied. One obvious explanation is that race might be more relevant or salient than university membership (Cunningham, 1986; Krebs, 1991). In a similar vein, Cosmides (Cosmides et al., 2003) noted that racial group membership defines coalitions and alliances during evolution and thus results in strong modulation of the neural substrates of emotional components of empathy.

Thus, relevant researches and my data suggest that relevant group membership might play a significant role in empathetic induction. But how can people rate differences of race? And, what does really means being members of two different races?

Since it is very difficult to directly measure the degree of genetic similarity with others, Krebs (Krebs, 1991) has suggested that one must rely on discernible cues to make such judgments. Clearly, racial indicators (e.g., skin color, hair texture) would seem to qualify as powerful cues of kinship and genetic similarity.

Once induced, empathy has been shown to have a substantial impact on subsequent behavioral and judgmental processes. For example, previous research has shown that inducing empathy for a target can lead to greater willingness to help that target (Batson, Klein, Highberger, & Shaw, 1995). In the same research Batson and colleagues demonstrated that empathetic induction led to a violation of participants' sense of morality: participants were asked to allocate resources to a number of individuals. Data showed that participants who were not induced to feel empathy made their allocation decisions in accordance with principles of justice and fairness. On the other hand, participants who were induced to feel empathy were far more likely to give preferential treatment to the target for whom empathy was felt even though they acknowledged that it was less fair and less moral. Batson and colleagues in 1997 (C. D. Batson et al., 1997) demonstrated that empathetic induction can affect attitudes toward members of stigmatized groups. More specifically, they found that inducing empathy led to more positive attitudes toward a young woman with AIDS. Furthermore recent findings showed that racial biases affect clinical pain management: Pletcher and colleagues (Pletcher, Kertesz, Kohn, & Gonzales, 2008) provided evidence that physicians withhold opioid treatment from Hispanic, Black, and Asian patients compared to

White patients, despite similar pain severity. The authors also noted this therapeutic disparity cannot be attributed to patient histories of alcohol and drug abuse as disproportionate treatment was most apparent in patients under the age of 12.

At the same time in United States of America there is extensive anecdotal evidence suggesting that a criminal predisposition is commonly attributed to Blacks. For example Johnson and colleagues (Johnson et al., 2002) reported that a number of Black leaders have expressed concern regarding the police practice of detaining and questioning individuals who meet the “profile” of criminals. These leaders contend that such profiling has led to a disproportionate number of arrests of minorities (“Endangered Species,” 1999). The American Civil Liberties Union became involved in the debate when they found that 75% of drug searches on the Maryland portion of Interstate 95 were of Black drivers, although they represented only 16% of the drivers on the road (“State Police Accused,” 1996). In addition, a report in the New York Times (“On Wealthy Island,” 1998) indicated that Blacks on Mercer Island, an affluent community in Washington, DC, have made numerous complaints about unjustified “traffic stops and harassment” from police.

Clearly, relationship between race empathy has many significant implications for real-world issues. Would racial differences decrease the probability that a Black teacher could empathize with a White student? Would racial similarity increase the probability that a White supervisor could empathize with a White subordinate? How a manager choose the person who has to be fired between a Black and a White worker? Although the teacher, the supervisor and the manager might make an effort to empathize, they are likely to be biased and experience less empathy for Blacks involved in their choices.

Taking together findings I discussed in the present doctoral thesis indicate that even if we are not aware, our body and our mind use internalized knowledge to address reactions and activities they engage to deal with social and physical world. My findings are coherent with Fisk's data and theories (Fiske et al., 2002) that predict people perceive outgroups on the base on warmth and competence. The so called Stereotyped Content Model predicts that perceived outgroups' warmth and competence affect what emotions outgroups elicit in perceivers. My data are consistent with this model and findings I reported seem to place race beside warmth and competence as a predictor of shared emotions.

At the same time path analysis models I discussed, for the first time seem to indicate that patterns of functional activations during a pain perception task are affected by internalized stereotyped associations. My findings show that the IAT scores significantly interact with the activation of right S1 area and they predict reactions of insula to social cues. My findings for the first time indicate that internalized associations affect not just brain activities but also how our brain elaborates external stimuli.

From my point of view, one of the most important elements is that my findings clearly show that we use fake or even false information to reason on the world. Like men in the Plato's cave, we are very likely to think that shadows of objects are the objects, shadows of people are people, stereotyped traits are real persons' traits. What it must be becomes what it is. Socially made stereotypes become more real than facts and stronger than empirical evidences. Stereotypes help us to create meanings in the world, like paths in a forest and force us to look for consistent cues, avoiding or even neglecting inconsistent facts.

My researches shed light on these phenomena and I think they provide useful knowledge to further understand relations between our brain, our mind and our behavior.

## References

- Adorno, T. V., Frenkel-Brunswik, E., Levinson, D. J., & Sanford, R. N. (1950). *The authoritarian personality*. New York: Harper.
- Aiken, L. S., & West, S. G. (1991). *Multiple Regression: Testing and interpreting interactions*. Thousand Oaks: Sage.
- Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley.
- Aron, A., & Aron, E. N. (1986). *love and the expansion of self*. Washington, DC: Hemisphere.
- Ashmore, R. D. (1981). Sex Stereotype and implicit personality theory. In D. L. Hamilton (Ed.), *Cognitive processes in stereotype and intergroup behavior*. Hillsdale, NJ: Erlbaum.
- Avenanti, A., Buetti, D., Galati, G., & Aglioti, S. M. (2005). Transcranial magnetic stimulation highlights the sensorimotor side of empathy for pain. *Nat Neurosci*, 8(7), 955-960.
- Avenanti, A., Minio-Paluello, I., Bufalari, I., & Aglioti, S. M. (2009). The pain of a model in the personality of an onlooker: Influence of state-reactivity and personality traits on embodied empathy for pain. *NeuroImage*, 44(1), 275-283.
- Avenanti, A., Paluello, I. M., Bufalari, I., & Aglioti, S. M. (2006). Stimulus-driven modulation of motor-evoked potentials during observation of others' pain. *Neuroimage*, 32(1), 316-324.
- Avenanti, A., Sirigu, A., & Aglioti, S. M. (2010). Racial Bias Reduces Empathic Sensorimotor Resonance with Other-Race Pain. *Current Biology*, 20(11), 1018-1022.
- Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal Autism Dev. Disord.*, 34, 163-175.
- Batson, C. (2005). Similarity and Nurturance: Two Possible Sources of Empathy for Strangers. *Basic and applied social psychology*, 27(1), 15-25.
- Batson, C., Early, S., & Salvarani, G. (1997). Perspective taking: Imagining how another feels versus imagining how you would feel. *Personality and Social Psychology Bulletin*, 23, 751-758.

- Batson, C. D. (1991). *The altruism question: Toward a social-psychological answer*. Hillsdale, NJ: Erlbaum.
- Batson, C. D., Chang, J., Orr, R., & Rowland, J. (2002). Empathy, attitudes and action: Can feeling for a member of a stigmatized group motivate one to help the group. *Personality & social psychology bulletin*, 28(12), 1656-1666.
- Batson, C. D., Duncan, B. D., Ackerman, P., & Buckley, T. (1981). Is empathic emotion a source of altruistic motivation? *Journal of personality and social psychology*, 40(2), 290-302.
- Batson, C. D., Klein, T. R., Highberger, L., & Shaw, L. L. (1995). Immorality from empathy-induced altruism: When compassion and justice conflict. *Journal of Personality and Social Psychology*, 68, 1042- 1054.
- Batson, C. D., Polycarpou, M. P., Harmon-Jones, E., Imhoff, H. J., Mitchener, E. C., Bednar, L. L., Klein, T. R., & Highberger, L. (1997). Empathy and attitudes: Can feeling for a member of a stigmatized group improve feelings toward the group? *Journal of Personality and Social Psychology*, 71(1), 105-118.
- Batson, D., Sager, K., Garst, E., Kang, M., Kostia, R., & Karen, D. (1997). Is empathy-induced helping due to self-other merging? *Journal of Personality and Social Psychology*, 73(3), 495-509.
- Brown, L. M., Bradley, M. M., & Lang, P. J. (2006). Affective reactions to pictures of ingroup and outgroup members. *Biol Psychol*, 71(3), 303-311.
- Bufalari, I., Aprile, T., Avenanti, A., Di Russo, F., & Aglioti, S. M. (2007). Empathy for pain and touch in the human somatosensory cortex. *Cereb Cortex*, 17(11), 2553-2561.
- Bushnell, M. C., Duncan, G. H., Hofbauer, R. K., Ha, B., Chen, J. I., & Carrier, B. (1999). Pain perception: Is there a role for primary somatosensory cortex? *Proceedings of the National Academy of Sciences of the United States of America*, 96(14), 7705-7709.
- Carr, L., Iacoboni, M., Dubeau, M.-C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences of the United States of America*, 100(9), 5497-5502.
- Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*. New York, NY: Guilford.



- Chauhan, B., Mathias, C. J., & Critchley, H. D. (2008). Autonomic contributions to empathy: Evidence from patients with primary autonomic failure. *Autonomic Neuroscience*, *140*(1-2), 96-100.
- Cheng, Y., Lin, C.-P., Liu, H.-L., Hsu, Y.-Y., Lim, K.-E., Hung, D., & Decety, J. (2007). Expertise Modulates the Perception of Pain in Others. *Current Biology*, *17*(19), 1708-1713.
- Cheng, Y., Yang, C.-Y., Lin, C.-P., Lee, P.-L., & Decety, J. (2008). The perception of pain in others suppresses somatosensory oscillations: A magnetoencephalography study. *NeuroImage*, *40*(4), 1833-1840.
- Cialdini, R. B., Brown, S. L., Lewis, B. P., Luce, C., & Neuberg, S. L. (1997). Reinterpreting the empathy-altruism relationship: When one into one equals oneness. *Journal of Personality and Social Psychology*, *73*(3), 481-494.
- Cosmides, L., Tooby, J., & Kurzban, R. (2003). Perceptions of race. *Trends in Cognitive Sciences*, *7*(4), 173-179.
- Craig, A. D., Reiman, E. M., Evans, A., & Bushnell, M. C. (1996). Functional imaging of an illusion of pain. *Nature*, *384*(6606), 258-260.
- Cunningham, M. (1986). Levites and brother's keepers: A sociobiological perspective on prosocial behavior. *Humboldt Journal of Social Relations*, *13*.
- Cunningham, W. A., Johnson, M. K., Raye, C. L., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2004). Separable Neural Components in the Processing of Black and White Faces. *Psychological science*, *15*(12), 806-813.
- Dale, P., M., B. E., Roberto, C., A., H. S., S., L. K., L., P. M., & G., W. M. (2008). *Principles of Cognitive Neuroscience*. Sunderland, MA, USA: Sinauer Associates Inc.
- Davis, K. D., Taylor, S. J., Crawley, A. P., Wood, M. L., & Mikulis, D. J. (1997). Functional MRI of Pain- and Attention-Related Activations in the Human Cingulate Cortex. *J Neurophysiol*, *77*(6), 3370-3380.
- Davis, M. H. (1980). Multidimensional approach to individual differences in empathy. *Journal Supplement Abstract Service (JSAS) Catalog of Selected Documents in Psychology*, *10*(85).
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*(1), 113-126.

- De Houwer, J. (In press). How do people evaluate objects? *Social and Personality Psychology Compass*.
- De Vignemont, F., & Singer, T. (2006). The empathic brain: how, when and why? *Trends Cogn Sci*, 10(10), 435-441.
- Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behav Cogn Neurosci Rev*, 3(2), 71-100.
- Decety, J., & Lamm, C. (2006). Human Empathy Through the Lens of Social Neuroscience. *TheScientificWorldJOURNAL*, 6, 1146-1163.
- Derbyshire, S. (2000). Exploring the pain “neuromatrix”. *Current Pain and Headache Reports*, 4(6), 467-477.
- Derbyshire, S. W. G., Jones, A. K. P., Gyulai, F., Clark, S., Townsend, D., & Firestone, L. L. (1997). Pain processing during three levels of noxious stimulation produces differential patterns of central activity. *Pain*, 73(3), 431-445.
- Deutsch, R., & Strack, F. (2005). Reflective and impulsive determinants of addictive behaviors. In R. W. Wiers & A. W. Stacy (Eds.), *Handbook of implicit cognition and addiction*. Thousand Oaks, CA: Sage.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *JPSP*, 56, 5-18.
- Devinsky, O., Morrell, M. J., & Vogt, B. A. (1995). Contributions of anterior cingulate cortex to behaviour. *Brain*, 118(1), 279-306.
- Eisenberg, N., & Carlo, G. (1995). Prosocial Development in Late Adolescence: A Longitudinal Study. *Child Development*, 66(4), 1179-1197.
- Eisenberg, N., & Morris, A. S. (2001). The Origins and Social Significance of Empathy-Related Responding. A Review of Empathy and Moral Development: Implications for Caring and Justice by M. L. Hoffman. *Social Justice Research*, 14(1), 95-120.
- Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 75-109). New York: Academic Press.
- Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion*, 15, 115-141.

- Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *J Pers Soc Psychol*, 82(6), 878-902.
- Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.
- Friston, K. J., Holmes, A. P., Poline, J. B., Grasby, P. J., Williams, S. C. R., Frackowiak, R. S. J., & Turner, R. (1995). Analysis of fMRI Time-Series Revisited. *NeuroImage*, 2(1), 45-53.
- Friston, K. J., Stephan, K. E., Lund, T. E., Morcom, A., & Kiebel, S. (2005). Mixed-effects and fMRI studies. *NeuroImage*, 24(1), 244-252.
- Fujiwara, H., Shimizu, M., Hirao, K., Miyata, J., Namiki, C., Sawamoto, N., Fukuyama, H., Hayashi, T., & Murai, T. (2008). Female specific anterior cingulate abnormality and its association with empathic disability in schizophrenia. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 32(7), 1728-1734.
- Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, 132, 692-731.
- Greenwald, A., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464-1480.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102, 4-27.
- Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, 109(1), 3-25.
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85, 197-216.
- Gu, X., & Han, S. (2007). Attention and reality constraints on the neural processes of empathy for pain. *NeuroImage*, 36(1), 256-267.
- Hamilton, D. L. (1981). *Cognitive processing in stereotyping and intergroup behavior*. Hillsdale, NJ: Erlbaum.
- Harris, L. T., & Fiske, S. T. (2006). Dehumanizing the lowest of the low: neuroimaging responses to extreme out-groups. *Psychol Sci*, 17(10), 847-853.

- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Emotional Contagion*. Cambridge, UK: Cambridge University Press.
- Hein, G., & Singer, T. (2008). I feel how you feel but not always: the empathic brain and its modulation. *Current Opinion in Neurobiology*, *18*(2), 153-158.
- Hornstein, H. A. (1978). Promotive tension and prosocial behavior: A Lewinian analysis. In L. Wispé (Ed.), *Altruism, sympathy, and helping: Psychological and sociological principles*. (pp. 177-207). New York: Academic Press.
- Hutchison, W. D., Davis, K. D., Lozano, A. M., Tasker, R. R., & Dostrovsky, J. O. (1999). Pain-related neurons in the human cingulate cortex. *Nat Neurosci*, *2*(5), 403-405.
- Jackson, P. L., Brunet, E., Meltzoff, A. N., & Decety, J. (2006). Empathy examined through the neural mechanisms involved in imagining how I feel versus how you feel pain. *Neuropsychologia*, *44*(5), 752-761.
- Jackson, P. L., Meltzoff, A. N., & Decety, J. (2005). How do we perceive the pain of others? A window into the neural processes involved in empathy. *Neuroimage*, *24*(3), 771-779.
- Johnson, J., Simmons, C., Jordan, A., MacLean, L., Taddei, J., & Thomas, D. (2002). Rodney King and O. J. revisited: The impact of race and defendant empathy induction on judicial decisions. *J Appl Soc Psychol*, *32*, 1208 – 1223.
- Keysers, C., Kaas, J. H., & Gazzola, V. (2010). Somatosensation in social perception. *Nat Rev Neurosci*, *11*(6), 417-428.
- Korol, B., & Kane, R. E. (1978). An examination of the relationship between race, skin color and a series of autonomic nervous system measures. *Pavlovian Journal of Biological Science*, *13*, 121–132.
- Krebs, D. (1975). Empathy and Altruism. *Journal of Personality and Social Psychology*, *32*(6), 1134-1146.
- Krebs, D. L. (1991). Altruism and egoism: A false dichotomy. *Psychological Inquiry*, *2*, 137-139.
- Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. Chicago: The University of Chicago Press.
- Lamm, C., Batson, C. D., & Decety, J. (2007). The Neural Substrate of Human Empathy: Effects of Perspective-taking and Cognitive Appraisal. *Journal of Cognitive Neuroscience*, *19*(1), 42-58.

- Lamm, C., Meltzoff, A. N., & Decety, J. (2010). How do we empathize with someone who is not like us? A functional magnetic resonance imaging study. *J. Cogn. Neurosci*, 22, 362 – 376.
- Lamm, C., Nusbaum, H. C., Meltzoff, A. N., & Decety, J. (2007). What Are You Feeling? Using Functional Magnetic Resonance Imaging to Assess the Modulation of Sensory and Affective Responses during Empathy for Pain. *PLoS ONE*, 2(12), e1292.
- Lamm, C., Porges, E. C., Cacioppo, J. T., & Decety, J. (2008). Perspective taking is associated with specific facial responses during empathy for pain. *Brain Research*, 1227, 153-161.
- Ledowski, T., Bromilow, J., Paech, M. J., Storm, H., Hacking, R., & Schug, S. A. (2006). Monitoring of skin conductance to assess postoperative pain intensity. *British Journal of Anaesthesia*, 97(6), 862-865.
- Levenson, R. W., & Ruef, A. M. (1992). Empathy: a physiological substrate. *J Pers Soc Psychol*, 63(2), 234-246.
- Lieberman, M. D. (2003). Reflexive and reflective judgment processes: A social cognitive neuroscience approach. In J. P. Forgas & K. D. Williams & W. von Hippel (Eds.), *Social judgments: Implicit and explicit processes*. New York: Cambridge University Press.
- Lieberman, M. D., Gaunt, R., Gilbert, D. T., & Trope, Y. (2002). Reflexion and reflection: A social cognitive neuroscience approach to attributional inference. *Advances in Experimental Social Psychology*, 34, 199-249. *Advances in Experimental Social Psychology*, 34, 199-249.
- Maner, J. K., Luce, C. L., Neuberg, S. L., Cialdini, R. B., Brown, S., & Sagarin, B. J. (2002). The Effects of Perspective Taking on Motivations for Helping: Still No Evidence for Altruism. *Personality and Social Psychology Bulletin*, 28(11), 1601-1610.
- Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality*, 40, 525-543.
- Mertz, H., Morgan, V., Tanner, G., Pickens, D., Price, R., Shyr, Y., & Kessler, R. (2000). Regional cerebral activation in irritable bowel syndrome and control subjects with painful and nonpainful rectal distention. *Gastroenterology*, 118(5), 842-848.
- Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2006). Dissociable Medial Prefrontal Contributions to Judgments of Similar and Dissimilar Others. *Neuron*, 50(4), 655-663.

- Morris, J. S., Ohman, A., & Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature*, *393*, 467-470.
- Morrison, I., & Downing, P. E. (2007). Organization of felt and seen pain responses in anterior cingulate cortex. *NeuroImage*, *37*(2), 642-651.
- Morrison, I., Lloyd, D., Di Pellegrino, G., & Roberts, N. (2004). Vicarious responses to pain in anterior cingulate cortex: Is empathy a multisensory issue? *Cognitive, Affective, & Behavioral Neuroscience*, *4*, 270-278.
- Morrison, I., Peelen, M. V., & Downing, P. E. (2007). The Sight of Others' Pain Modulates Motor Processing in Human Cingulate Cortex. *Cereb. Cortex*, *17*(9), 2214-2222.
- N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, Bernard Mazoyer, & M. Joliot. (2002). Automated Anatomical Labeling of activations in SPM using a Macroscopic Anatomical Parcellation of the MNI MRI single-subject brain. *NeuroImage*, *15*(1), 273-289.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231-259.
- Ogino, Y., Nemoto, H., Inui, K., Saito, S., Kakigi, R., & Goto, F. (2007). Inner Experience of Pain: Imagination of Pain While Viewing Images Showing Painful Events Forms Subjective Pain Representation in Human Brain. *Cereb. Cortex*, *17*(5), 1139-1146.
- Olsson, A., & Ochsner, K. N. (2008). The role of social cognition in emotion. *Trends in Cognitive Sciences*, *12*(2), 65-71.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, *46*, 598-609.
- Pettigrew, T. F., & Meertens, R. W. (1995). Subtle and blatant prejudice in western Europe. *European Journal of social psychology*, *25*, 57-75.
- Peyron, R., Garcia-Larrea, L., Gregoire, M.-C., Costes, N., Convers, P., Lavenne, F., Mauguier, F., Michel, D., & Laurent, B. (1999). Haemodynamic brain responses to acute pain in humans: Sensory and attentional networks. *Brain*, *122*(9), 1765-1780.
- Peyron, R., Laurent, B., & García-Larrea, L. (2000). Functional imaging of brain responses to pain. A review and meta-analysis (2000). *Neurophysiologie Clinique/Clinical Neurophysiology*, *30*(5), 263-288.
- Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2000). Performance on indirect measures

of race evaluation predicts amygdala activation. *Journal of cognitive neuroscience*, 12(5), 729-738.

Plato. (514-520 b.c.). *The Republic. Book VII*.

Pletcher, M. J., Kertesz, S. G., Kohn, M. A., & Gonzales, R. (2008). Trends in Opioid Prescribing by Race/Ethnicity for Patients Seeking Care in US Emergency Departments. *JAMA: The Journal of the American Medical Association*, 299(1), 70-78.

Preston, S. D., & de Waal, F. B. (2002). Empathy: Its ultimate and proximate bases. *Behav Brain Sci*, 25(1), 1-20; discussion 20-71.

Preston, S. D., & de Waal, F. B. M. (2001). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, 25(01), 1-20.

Rae Westbury, H., & Neumann, D. L. (2008). Empathy-related responses to moving film stimuli depicting human and non-human animal targets in negative circumstances. *Biol Psychol*, 78(1), 66-74.

Rorden C, & Brett M. (2000). Stereotaxic display of brain lesions. *Behav Neurol*, 12, 191-200.

Russell, K. Y., Wilson, M., & Hall, R. E. (1992). *The color complex: The politics of skin color among African Americans*. New York: Harcourt Brace Jovanovich.

Saarela, M. V., Hlushchuk, Y., Williams, A. C. d. C., Schurmann, M., Kalso, E., & Hari, R. (2007). The Compassionate Brain: Humans Detect Intensity of Pain from Another's Face. *Cereb. Cortex*, 17(1), 230-237.

Santo, J. L., Arias, L. M., Barolat, G., Schwartzman, R. J., & Grossman, K. (1990). Bilateral cingulumotomy in the treatment of reflex sympathetic dystrophy. *Pain*, 41(1), 55-59.

Singer, T., & Lamm, C. (2009). The Social Neuroscience of Empathy. *Annals of the New York Academy of Sciences*, 1156(The Year in Cognitive Neuroscience 2009), 81-96.

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

Singer, T., Seymour, B., O'Doherty, J. P., Stephan, K. E., Dolan, R. J., & Frith, C. D. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature*, 439(7075), 466-469.

- Smith, E. R., & DeCoster, J. (2000). Dual process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, 4, 108-131.
- Strack, F., & Deutsch, R. (2004). Reflective and Impulsive Determinants of Social Behavior. *Personality and Social Psychology Review*, 8(3), 220-247.
- Strack, F., Werth, L., & Deutsch, R. (2006). Reflective and impulsive determinants of consumer behavior. *Journal of Consumer Psychology*, 16, 205-216.
- Svensson, P., Minoshima, S., Beydoun, A., Morrow, T. J., & Casey, K. L. (1997). Cerebral Processing of Acute Skin and Muscle Pain in Humans. *J Neurophysiol*, 78(1), 450-460.
- Tajfel, H. (1981). *Human groups and social categories: studies in social psychology*. Cambridge: Cambridge University Press.
- Tajfel, H., & Turner, J.C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (pp. 7-24). Chicago: Nelson-Hall.
- Tajfel, H., Flament, C., Billig, M. G., & Bundy, R. P. (1971). Social categorization and intergroup behavior. *European Journal of social psychology*(1), 149-178.
- Talbot JD, Marrett S, Evans AC, Meyer E, Bushnell MC, & GH., D. (1991). Multiple representations of pain in human cerebral cortex. *Science*, 251, 1355-1358.
- Talbot, J. D., Marrett, S., Evans, A. C., Meyer, E., Bushnell, M. C., & Duncan, G. H. (1991). Multiple representations of pain in human cerebral cortex. *Science*, 251(4999), 1355-1358.
- Teige-Mocigemba, S., Klauer, C., & Sherman, J. (in press). A Practical Guide to Implicit Association Tests and Related Tasks. In B. Gawronski & B. K. Payne (Eds.), *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications*. New York: Guilford Press.
- Teige-Mocigemba, S., Klauer, K. C., & Sherman, J. W. (2010). A Practical Guide to Implicit Association Tests and Related Tasks, *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications*. New York: The Guilford Press
- Wegner, D. M. (1980). The self in prosocial action. In D. M. W. R. R. Vallacher (Ed.), *The self in social psychology* (pp. 131-157). New York: Oxford University Press.



- Wicker, B., Keysers, C., Plailly, J., Royet, J.-P., Gallese, V., & Rizzolatti, G. (2003). Both of Us Disgusted in My Insula: The Common Neural Basis of Seeing and Feeling Disgust. *Neuron*, *40*(3), 655-664.
- Wilder, D. A. (1981). Perceiving persons as a group: categorizations and intergroup relations. In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior*. Hillsdale, NJ: Erlbaum.
- Xu, X., Zuo, X., Wang, X., & Han, S. (2009). Do You Feel My Pain? Racial Group Membership Modulates Empathic Neural Responses. *J. Neurosci.*, *29*(26), 8525-8529.