

Framework Speech
Sunday July 18, 2004
2. Is Our Brain Ethical?
The Neurobiology of Moral Thought

***BENUMBING AND MORAL EXALTATION IN DEADLY MARTYRS: A
VIEW FROM NEUROSCIENCE*** (*)

Adolf Tobeña
Department of Psychiatry and Forensic Medicine, Institute of
Neuroscience, Universitat Autònoma de Barcelona

(*) This paper is based on Tobeña, A. (2004) *Mártires mortíferos: biología del altruismo letal* (Deadly Martyrs: The Biology Of Lethal Altruism), Barcelona: Tusquets Ed. (in press).

BENUMBING AND MORAL EXALTATION IN DEADLY MARTYRS: A VIEW FROM NEUROSCIENCE

“Of all the differences between man and the lower animals, the moral sense conscience is by far the most important. This sense has a rightful supremacy over every other principle of human action; it is summed up in that short imperious word ought, so full of high significance. It is the most noble of all attributes of man, leading him without a moment's hesitation to risk his life that of a fellow creature; or after due deliberation, impelled simply by the feeling of right or duty, to sacrifice it in some great cause.”

Charles Darwin (1871): *The Descent of Man and Selection in Relation to Sex*

Suicide attacks, through their unpredictability and devastating potential have become one of the most disturbing weapons of our day. Since the tremendously lethal nature was demonstrated in the attacks of 09-11-01 on New York and Washington, there has been no shortage of volunteers eager to emulate such deeds — as was only to be expected once it became clear that the North American giant could be wounded. Indeed, the tactic of suicide attacks has since become widespread, and all too common in several of the world's hotspots. Moreover, despite the efforts invested in preventing this type of attack, they have also — occasionally but with ominous consequences — been carried out in more peaceful parts of the planet. Nowadays, the possibility that weapons of mass destruction might fall into the hands of an organization capable of harboring suicide bombers is a major concern for security services around the world. We live under this threat that, in the end, hangs on the determination and nerve of a handful of individuals: the combatative cell that provides a home for the phenomenon of suicide terrorism (4, 54). Given that suicide attacks depend ultimately, on conviction — that is, on events that take place in the brain of the individual — it may be useful to consider such behavior from the perspective of neuroscience. It should be remembered that suicide for non-combatative purposes has been and continues to be a compulsory area of study for clinicians although this form of self-destruction has little to do with the behavior of intelligent beings here. Furthermore, recent advances in the neurobiology of morality may help shed some light on the exceptional behavior of these deadly martyrs.

Among the most bewildering aspects of the behavior of deadly martyrs there is one which is particularly striking, namely, the strange moral boundaries

established by those who carry out suicide attacks. These boundaries are usually shared with members of their cell or group and with a certain sector of society in which the phenomenon is bred (4, 31). Kamikazes do not only kill and carry out horrific massacres with the utter conviction that they are doing good, but also act in accordance with values and beliefs associated with what they regard as the greatest good (26). They kill in the name of supreme justice and duty. They exterminate in obedience of infinite and impartial Providence having given themselves up to the orders of their just and discerning guide. Deadly martyrs are the most spectacular examples of an impregnable moral code towards one's own group which co-exists, without any apparent contradiction, alongside a radical amorality towards members of another group. The intergroup barrier clearly delimits the domain of moral behavior. Hence, in their own communities suicide killers are regarded as heroes and hailed as paragons of virtue, while among those who are suffering from having been chosen as a target they are held up as cold-hearted butchers.

It has often been said that treating the "other" as an object is a necessary condition for this intimate relationship between deadliness and virtue to be established (39, 43, 52). The quashing of all empathy with the suffering adversaries finds favorable ground in military confrontations, and it undoubtedly formed part of the long history of deadly hostilities between human groups in conflict. Moreover, in the age of smart weapons that enable selective strikes to be made with often exquisite accuracy, it remains part of the tradition of frontline forces and, more often than not, the indoctrinating propaganda of the populations who stand behind them. In order to be effective in combat the closer one gets to such predatory reactions the better. When one's opponent is seen as vermin to be captured or exterminated, empathy or other inhibitions induced by moral learning to respect the other count for nothing. In the case of these martyrs who die in an attack, however, an additional step is taken in this paradoxical combination of moral and amoral tendencies, namely, a total disdain for existence itself — irremissible squandering of individual interests in the pursuit of communal justice. Such waste has lately taken on an obscene form in the explosive self-dismemberment of suicide bombers. The immediate family of the martyr is not thus to be ennobled and have honors or advantages bestowed upon them, yet one should be borne in mind that those who sacrifice themselves have no family. It is not guaranteed that this will be the case, even when they act under contract. The certain thing is the contribution without recompense to group interests. Consequently, we are faced with a particularly challenging area of enquiry: moral convictions.

The Origins of Morality

In addressing this issue from a biological perspective it is necessary begin by referring to Darwin. For him it was clear that human morality is usually restricted to the confines of the group (13). Despite being the author of the eloquent statement with which I began this text he did not have the time to moderate the enthusiasm characteristic of his Victorian age. A few pages later in the same essay he offers a host of examples through which he proclaims insistently that the moral sense of humans is not extended to all their fellows and that tribal limits usually mark out the territory of confrontation where ties of sympathy, affection and the exchange of favors which sustain moral inclinations and obligations are rudely extinguished. Darwin was the first to discern that frequent wars between neighboring groups could induce natural selection of the most demanding loyalties in many individuals — a conjecture that has regained favor in recent decades despite having been overlooked for most of the 20th century (2, 22, 28, 48, 54, 64).

Yet long before he addressed this issue Darwin had already sought to bring down the lofty pretensions of that most moral of animals, proposing connections with the behavior of other social animals and putting forward various conjectures as to why the peculiar obligations towards our fellows that we call morality should arise. In his words:

“Any animal whatever, endowed with well-marked social instincts, the parental and filial affections being here included, would inevitably acquire a moral sense or conscience, as soon as its intellectual powers had become as well, or nearly as well developed, as in man. For, firstly, the social instincts lead an animal to take pleasure in the society of its fellows, to feel a certain amount of sympathy with them, and to perform various services for them [...]. Secondly, as soon as the mental faculties had become highly developed, images of all past actions and motives would be incessantly passing through the brain of each individual [...]. Thirdly, after the power of language had been acquired, and the wishes of the community could be expressed, the common opinion how each member ought to act for the public good would naturally become in a paramount degree the guide to action [...]. Lastly, habit in the individual would ultimately play a very important part in guiding the conduct of each member; for the social instinct, together with sympathy, is, like any other instinct, greatly strengthened by habit and so consequently would be obedience to the wishes and judgment of the community.”

This passage provides us with a well-developed working hypothesis which to enquire into the origins of moral behavior. Darwin placed foundations of human morality on the following equation: sociability recognition of emotions in fellows + regular exchange of services + obedience to the wishes of the community. All of this being dependent on a brain that developed enough to make such attributes viable. In other words, that there is sufficient neurocomputational power to enable these mechanisms to work and achieve the necessary subtle balances during private self-scrutiny (conscience) and those which we perceive in the form of scruples, inclinations and obligations. Table I shows the pioneering conjectures of Darwin alongside the mechanisms of biological selection that have since been demonstrated, both in terms of ultimate functionality (adaptive strategies) and the ordinary operations that support it (behavior and its neural mechanisms). It can be seen that Darwin's view was basically correct. It should be borne in mind that the aim of convincingly elucidating the origins of altruism became a fundamental problem in evolutionary thought — for it was necessary to explain the genesis of cooperative behavior at the expense of individual interests, including pro-group sacrifice. Although debate about most of the mechanisms shown in Table I tended to peter out as new evidence has been gathered, controversy continues to rage over the viability of pro-group altruism (8, 14, 19, 43, 44, 62, 65).

TABLE I

HUMAN MORAL BEHAVIOR

Definition: Behavior, attitudes, desires and beliefs that take into account the interests of fellow humans.

“In animals with social instincts, and whose intellectual powers are developed enough to enable them to contemplate the result of interactions with their fellows, cooperative tendencies and moral conscience should come about through selection.”

Charles Darwin (1871): *The Descent of Man and Selection in Relation to Sex*

<i>Darwinian Conjectures</i>	<i>Biological Selection</i>
<u>Origin of Moral Conscience</u>	<u>Demonstrated Mechanisms</u>
Parental/filial affection	Nepotism (kin selection)
Exchange of services/fulfilling obligations	Direct reciprocity (reciprocal altruism)
Reputation in the group/community	Indirect reciprocity
Obedience to social norms	Altruism induced by punishment
Loyalty/pro-group sacrifice	Group selection
Abstract moral principles	???

Neural Pre-Requisites

Recognition of fellows and markers of identity
Detection/expression of social emotions (affection, empathy, guilt, etc.)
Operational memory for comparing/weighing up and predicting the results of social interactions
Traits of temperament (loyalty, obedience, credulity, etc.)

The main body of results from studies of animal behavior, computer simulations of evolutionary contests and experiments with humans indicates forms of altruism which are non-nepotistic and not based on direct reciprocity do exist (11, 25, 40, 51, 53); there are indeed propensities toward established cooperative relationships that go beyond biological filiations, the exchange of favors and the weighing up of returns. Non-restricted altruism exists and is perpetuated because it is maintained by various mechanisms: the generation of reputation, intra-group size and signaling, moralizing punishment and unconditional trust (7, 17, 18, 32, 36, 45, 63). Altruism at a price and with guaranteed returns occurs because it tends to promote gains in biological yieldings, not only for the recipients of help but also for the givers. Of course the latter does not apply in cases (like our object of study here) where the givers perish too soon as a result of their exaggerated investments; yet these strategies of cooperation at a price do prove worthwhile for those givers who stay around.

So we must take into account the trait of altruism and be able to measure it reliably, individual differences in it (6, 12, 30). This should be our starting point. For the existence of this trait implies that among all populations there will be pockets of extreme altruists willing to make high cost investments of trust in others, investments that involve great personal risk. This is the supposition which, until recently, the prevailing models in biology and economics were determined to discard, for the simple reason that in all populations there are systematic egoists able to take advantage of the propensity toward innocence and sacrifice of genuine givers. Yet despite these rather unflattering views regarding the place of extravagant generosity in the process of evolution there are indeed mechanisms geared toward fostering and maintaining it (19, 34, 52).

A Neuroimage of Moral Decisions

When behavior is accompanied by an awareness of “duty” and “sacrifice” then we are faced with a specifically human form of certain attributes that have firm adaptive foundations and which reflect selective forces operating at different levels. This conscious state of moral inclinations was, for many years, ignored by neurobiological research, due, among other things, to a lack of procedures for carrying out reliable empirical investigations in the area. Therefore for too long the work of evolutionary biologists aimed at outlining the adaptive means required to explain variability in altruism remained disconnected from that of psychologists and neurobiologists who sought to describe the mechanisms behind behavior and social cognition. In recent years, however,

situation has taken a radical turn and the area of research known as the cognitive neuroscience of morality is now a hive of activity (1, 23). The first stimuli for this shift came from the renewed interest in neurological patients with selective deficits or anomalies in moral decision-making (3, 57), and from there attention switched to the exploration of pro-social and moral inclinations in many different areas, the research teams involved bringing together neurologists, radiologists, philosophers, engineers and other professionals from various fields with interest in these unsolved enigmas.

In recent decades several procedures have been designed for obtaining quantitative measures of moral decisions. One of the most useful methods for developing this kind of “experimental ethics” consists in posing moral dilemmas, that is, problems whose possible solutions affect the interests of other people and where subjects are asked to make a decision which they can not go back on. By introducing variations into these dilemmas, or by presenting different versions one after another, it is possible to analyze response trends and individual differences. As some of these problems are by their very nature simple, and can be presented in the form of brief texts on a screen, it is not surprising that the process of moral decision-making has become a topic of study within the field of neuroimaging. This type of enquiry was first carried out by a team of researchers from the University of Princeton (21), in a study comparing two dilemmas which have proved to be a headache for philosophers concerning with the origin of ethical values.

The first of these is known as the trolley dilemma, and can be formulated as follows: a runaway trolley will kill five people on the track ahead if it continues on its course. The only way to avoid this is by hitting a switch, either on the trolley or at the control center, that will turn the trolley onto a side track where it will kill one person who happens to be crossing at that moment. Nobody in the trolley will be injured in any way. The dilemma is therefore whether it is right to hit the switch and change the course of the trolley. Many people respond by saying that, under these circumstances, it is acceptable to sacrifice the life of one person to save the other five. It should be immediately clear that there are many possible variations on the dilemma: changing the number of people, changing the relationship to or familiarity with the people to be saved or sacrificed, and introducing different degrees of chance (instead of a passer-by crossing the road at just the wrong moment it could be a man working on the overhead power lines), among other variables able to add interest to the dilemma. The result which concerns us, however, is that in the basic situation people opt to sacrifice deliberately one person to save the lives of the other five about to be run over by the tram.

The second problem derives from the first and is known as the footbridge dilemma. It goes like this: a runaway trolley will kill five people before it reaches the buffers at the end of the line. There is no way of avoiding this. You are standing on a footbridge over the tracks and realize what will happen. You then notice that next to you there is a large stranger leaning over the rail in order to see what's going on below. If you push the stranger onto the tracks, killing him, his bulky body will stop the trolley from reaching the others, thus saving five lives. Nobody in the trolley will be injured in any way. Do you push him off the tracks and save five lives? Under these circumstances most people change their mind and consider that it is not right to use the life of one person to save five others. We are thus faced with a puzzling enigma: human beings tend to consider it acceptable to sacrifice deliberately the life of one person in exchange for five when all that is required is the flick of a switch, but not when it would involve pushing the actual person to be sacrificed. Given that the costs and benefits are the same in both situations, what produces this radical switch in most people's moral criterion? It seems that the evidence cannot be explained through recourse to ethical theories. The team of neurobiologists at Princeton hypothesized that the second dilemma leads to greater emotional involvement than the first due to the proximity of and contact with the victim, hence the change in most people's moral criterion. In order to corroborate this they decided to use functional magnetic resonance imaging (fMRI) to scan subjects while they were resolving the dilemmas.

Their first experiment involved nine university students (five men and four women) who gave their informed consent to take part in the sessions. They were each presented with sixty dilemmas which appeared on a screen that was visible from inside the fMRI tube. Each dilemma involved three screen stages: the statement, the question, and, on the final screen, the words "appropriate" or "inappropriate". All the dilemmas lasted a maximum of 45 s and reaction times were measured in each case, along with whole and regional brain activity during the decision-making period. The sixty dilemmas were presented at random and belonged to one of three conditions, twenty dilemmas in each: 1) personal/close (with greater emotional involvement), comprising variations of the footbridge dilemma and equivalents; 2) impersonal/distant (variations on the trolley dilemma and equivalents) with less emotional charge; and 3) non-moral dilemmas (harmless choices, such as taking the subway rather than the bus into town or having dinner at a Chinese rather than a Thai restaurant). The results indicated that Group 1 dilemmas (personal/close) significantly increased brain activity in areas associated with emotional regulation: the medial frontal gyrus, the posterior cingulate gyrus and the angular gyrus. In contrast, impersonal/distant dilemmas (Group 2) produced increased brain activity in areas concerned with

working memory, but with no appreciable increases in areas of emotic regulation. All these results were compared with baseline activation levels harmless dilemmas.

A second experiment carried out with nine new subjects produced identical data in terms of regional brain activation, thus lending further support to the findings. The researchers also studied the relationship between reaction times and the type of response to the dilemmas presented. In personal/choice dilemmas subjects found it much more difficult (a delay of around 1 s more to give the response “appropriate” rather than “inappropriate”, while the opposite trend was found with impersonal/distant dilemmas. In other words, the subjects who opted to give the fateful push in the footbridge dilemma took longer to reach this decision, as did those who considered it inappropriate to a switch and change the course of the trolley in the first dilemma. This lends support to the initial assumption that the emotional charge associated with the situation interferes with and modulates moral decisions. Through additional control studies using much more demanding dilemmas (involving the sacrifice of one’s own child) the authors were able to demonstrate that the results obtained were not due to majority or minority bias. This research therefore opens a door to the analysis of the neural mechanisms underlying moral value. In what follows, we will consider other ways in.

The Neurobiology of Cooperative Interactions

The Ultimatum Game is one of the experimental situations used by economists to study egoistic or altruistic choices in social interactions. It is a one-shot and non-negotiable game that consists essentially in a “take it or leave it” offer. The simplicity of the approach has led researchers to use neuroimaging techniques with subjects in order to monitor their brain activity while they are making decisions. Thus, work has begun on identifying the brain areas involved in decision-making regarding dilemmas of economic exchange (20).

In the Ultimatum Game one of the participants receives a sum of money which can be kept on one condition: he must offer a fraction of it to the other participant who he has just been introduced to prior to the experiment. Both players first (the Proposer) and the second player (the Responder) will have carefully read the rules of the game: these are simply that if the Responder freely accepts the Proposer’s offer then they both keep the share agreed. However, if the Responder rejects the offer, neither of them gets anything and the experimenter keeps the money. There is no negotiation or opportunity to change your mind, hence the name “Ultimatum Game.” Players are situated in separate cubicles

make their responses via a computer terminal. Several rounds are usually played but always with different subjects in order to respect the rules of the game. The supposed rationality that people use in optimizing costs and benefits and which economic theory is based would predict that Responders should accept any offer because something is always better than nothing. Likewise, Proposers assuming such rationality in the Responder, should propose a deal that is favorable to themselves as they alone decide how the money is to be shared out.

However, such studies of experimental economics have found that the assumptions are not borne out by reality: offers giving 20% of the total to Responders have a 50% chance of being rejected, while for those offering more than 20% the rejection rate rises sharply and may exceed 70%. In other words, a majority of Responders prefer to do without a cash gift if this prevents a stranger from keeping the lion's share in a deal regarded as unfair. Put another way, under these circumstances over 50% of people punish a fellow human being even though this means losing out themselves. This situation is therefore a glaring example of altruistic punishment (17): there is a palpable cost with no benefit other than having scuppered the plans of another who was deemed to be taking unfair advantage. Proposers, in reality, anticipate such behavior and tend to make offers close to an equal share; this has proved to be the case in many experiments carried out in different parts of the world, although there are some variations depending on culture and relationships to money. Feedback given to participants after the experiments suggests that it is the sense of indignation which leads to the altruistic punishment: any interest in making a quick buck is abandoned through the desire to reject a share regarded as unfair. Thus do the participants resolve a genuine dilemma where emotional reactions (spite, envy, scorn) modulate the opportunity for financial gain.

The neuroimaging studies consisted in exploring the brain activity of subjects who received "ultimatum" proposals from ten Proposers whom they had met briefly just prior to the brain scans being carried out (47). Once inside the fMRI scanner they received, via a screen and in a sequence that was identical for each round, the offers to be accepted or rejected by pushing a button situated within hand's reach. After 12 s of preparation the photograph and name of one of the ten Proposers appeared, for 6 s, on the screen. The proposed share then appeared and the Responder, from inside the magnet, had 6 s to reach a decision. Once the offer had been accepted or rejected the winnings of each player (or the amount of the gift) was shown on the screen. This procedure was repeated over consecutive rounds with the different Proposers, interspersed at random with both analogous offers made by a computer program (in this case the screen showed the image of a laptop acting as the Proposer) and control runs which consisted in trying to win the amount shown on the screen by responding

quickly as indicated, there being no offer or dilemma in these cases. Each scanned subject underwent thirty sequences in all, the hand-outs coming from an initial pot of \$10 per round provided by the experimenters. The accumulated amounts at the end of each session were sufficient recompense for participants.

In terms of the decisions made the results were consistent with previous research, and in this sense the brain scan equipment did not prove to be a constraint. Equal offers (\$5 for each player) were accepted without hesitation whatever their source, while those giving \$7 to the Proposer and \$3 to the Responder produced the excellent acceptance rate of 90%. However, offers of \$8 for the Proposer and \$2 for the Responder were only accepted half the time (50% rejection), while those giving \$9 to the Proposer and just \$1 to the Responder produced a rejection rate of 70%. These results confirm the tendency to punish unfair offers even when this means losing money. It is also interesting that a good number of Responders decided to apply an analogous criterion of altruistic punishment, albeit somewhat weaker, to some of the offers made by the computer: 20% of subjects decided against keeping the \$2 offered by the computer when the computer was going to keep \$8 for itself, and almost half chose to reject the single dollar if the computer kept \$9. These results confirm the unique power of the drift toward a sense of equal justice shown by many human beings; it should be noted that this drift is not solely a human prerogative, as some primates also reject unfair deals in circumstances where they obtain a reward for little or no effort (9). These rejections include decisions without tasty food when their companion in the interaction has been rewarded ostentatiously with something better.

Returning to the experiment, the brain areas showing increased activity in the face of unfair offers were the dorsolateral prefrontal cortex, the anterior cingulate gyrus and the anterior insula. Increased neural activity in the first two areas probably corresponds to the cognitive effort involved in weighing complex options (prefrontal circuits) in a conflictive dilemma situation (cingulate circuits): it should be noted that unfair offers meet these requirements because, when rejection is contemplated, they involve the loss of a gift or consideration for moral compensation. Previous studies of the functions of these brain areas support this assumption. However, the most interesting result without doubt is the bilateral activation of the anterior insula. Research has associated these areas with the processing of physical disgust (37), that is, unpleasant reactions to olfactory or taste stimuli that, ordinarily, people find disgusting or repugnant. If these results are confirmed in subsequent studies of economic games it will demonstrate that the neural circuits involved in producing feelings of disgust at the sight or consumption of food that has g

off, or in response to unpleasant body odors, are the same as those used to process feelings of indignation at the unfair behavior of another human being. This, of course, makes perfect sense in terms of saving neural resources. Studies such as these are opening up the field of neuroeconomics (10, 20), and one who has already produced strong evidence relating the development of trust between individuals in financial interactions to the selective activation of pleasure areas in the brain and the secretion of neurohormones that modulate responses such as attachment and affective proximity (44).

The Neuroimaging of Amorality

The possibility of working with normative individuals in order to study moral decision-making constitutes an important advance, as until recently research into the biology of human morality depended heavily on data obtained from people with mental disorders or those with lesions in specific brain areas. This long tradition of studies in people with selective brain lesions has nevertheless been important in providing useful clues. The combination of clinical and anatomical evidence from children, young people and adults with singularly conflictive behavior and measurable neural anomalies has enabled various brain systems to be identified as substrates for the crystallization of moral values (3, 57). Thus, the appearance of reasoning and normative moral reactions depends necessarily on certain brain areas and circuits being intact; if they are damaged in some way striking deficits in moral behavior, that are hard to remedy, arise — even though the ability to perceive the content, meaning and value of social conventions and ethical norms remains unimpaired. In other words, lesions to specific areas of the brain leave fully intact motor, sensory and perceptive functions, such as memory, language and most aspects of cognitive acuity, but produce deficits in or the absence of moral behavior.

These results are particularly evident in children who have suffered very early lesions in specific areas of the prefrontal cortex (specifically, in the rostral anterior, medial and basal regions of the cerebral hemispheres). However, when such lesions occur in an already developed brain patients conserve their memory of norms and may agree entirely with other people when asked about moral limits in various situations, even though this is often not borne out by their behavior. It should be concluded, therefore, that in these patients the learning of moral boundaries has been able to occur and take root, despite the fact that a subsequent lesion leads to repeated transgressions of such inhibitory boundaries. However, in the few children studied in similar conditions and over time, development into adults not even the awareness of norms seems to have been

properly implanted (3). Not only do they regularly transgress ordinary moral boundaries but also show both clear signs of never having learnt them and absence of guilt or remorse over their behavior. In any case, despite the power of these results it is likely that the type of lesion referred to here produces gross indirect alterations in the circuits responsible, in intact brains, for moral decision-making processes. The door is now open, however, to studies of ordinary subjects of the neural correlates of that particular set of attitudes and decisions which we term moral behavior — and, of course, to studies of abnormal behavior and propensities.

The area of emotions may again prove particularly relevant, just as we saw with the studies involving dilemmas. Ordinary people changed their moral criterion according to the degree of empathy invoked by the situation: in the case of the disaster that was about to be caused by the runaway trolley the emotional closeness or distance from the person to be sacrificed modified the reaction to the suffering of others and the subsequent choice of action. Empathy is a crucial emotion in defining moral and amoral reactions as it is part of the basis of human sociability. It should be remembered that this capacity to tune into the happiness or suffering of others was, for Darwin, the first requisite of well-founded social instincts. Furthermore, there is now a substantial body of research concerning the neural mechanisms underlying empathic attributes in both animals and humans (15, 38, 51). However, alongside empathy there is a wide range of other moral emotions (disdain, indignation, guilt, remorse, shame, etc.) that may be evoked by positive and negative interactions with others, by the obligations established during these interactions and by the contrast between expectations and actual outcomes (23) — emotions that are increasingly beginning to be addressed by researchers. It is reasonable to suppose that, like empathy, these other emotions also influence moral decisions, particularly in interactions that involve important commitments between those involved. Under such circumstances the role of pretense and lying is, of course, crucial, and neural-based methods for detecting this are already beginning to bear fruit (56).

Therefore, the future is likely to see a proliferation of studies that examine normative individuals in various frames of mind involving moral emotions and decisions at the same time as scanning their brain activity with sufficiently powerful neuroimaging and computer equipment. Some authors have used a procedure in which subjects are presented with vignettes showing conflictual interactions between different people (abusers and abused, or executioners and their victims) and then asked to evaluate them on specially-designed scales. The results of studies carried out with children and adolescents have already served to establish quantitative degrees of morality (5). Such approaches are easily applied even with neuroimaging equipment and some preliminary findings have

already been reported (33). However, research must also address situations more closely reflect real personal involvement and the conditions of every life because there is a sizeable body of data indicating there to be substantial differences between people's responses to abstract dilemmas and vignettes opposed to circumstances that touch upon their real lives (28). The degree of suffering caused by temporary social rejection in video games designed for entertainment purposes, in which various players take part, has also been studied using neuroimaging equipment (16), suggesting that even the experience of ostracism as a form of social punishment can be investigated using this apparatus. The door is wide open, therefore, and it is to be hoped that research in this field focuses particularly on adolescents and young people as it is they who are usually at the forefront of combative cells fostering harmful tactics.

Moral Deficiency and Fanaticism

We can now return to our objective of shedding light on the origins of deadly martyrs. The passionate belief in a doctrine widens the gap between groups, a gap which as a matter of course already serves to limit cooperative behavior and feelings of solidarity. This form of combative fervor, fanatic sectarianism, usually takes root within tiny pockets of society, but occasionally and with an all-encompassing vigor, it colors the beliefs of large communities. The sectarian creed governs so strongly inside the bubble of its associated doctrine that the propensities and decisions produced therein are often incomprehensible to those who contemplate the process from the outside. It can lead to a state of moral deficiency, for not only is irreparable damage done to those who remain outside the realm of "undisputed virtue" (the exogroup) but outrages are also committed against those suspected of dissidence within the confines of "virtue" — barbarities that may even come to affect members of the same family or the local neighborhood. Indoctrination can result in people's inclinations going beyond the most basic social instincts, the tendencies which serve to build the most peremptory links in the chain of morality (one thinks of recent cases of young-mother suicide bombers in Palestine whose actions irreparably condemn their own children to be orphans). The point is, therefore, that indoctrination can modify people's moral inclinations and obligations to such an extent that it leads to behavior patterns comparable to those seen in personality disorders.

The notion of "moral deficiency" underlies the first attempts to describe those singular variants on human temperament referred to as psychopathy (34). The core characteristic of this type of personality is precisely the striking lack

moral emotions. However, although psychopaths are peculiar people they do stand clearly apart from the profuse variety of human character. Indeed, the highly special temperament has yet to be associated with ostensible brain lesions, despite their having been subjected to detailed studies in the search for them. Much evidence has recently been reported, however, about possible disorders of functioning and connectivity in the same neural circuits that, when damaged, produce amoral behavior (27, 41). Although these data are clearly important they should not, at present, be regarded as definitive because the methods used to access these subtle aspects of brain functioning still lack the necessary degree of resolution. In contrast, studies of emotional reactions in individuals with a psychopathic temperament do provide consistent profiles. Such research suggests them to be people who are perfectly capable of developing intense desires and of feeling joy, boredom or rage, to give some examples of everyday emotions, but who find it difficult to experience genuine remorse, foolishness or shame in a genuine way — that is, they struggle with the shades of emotion normally associated with interactions in which someone, usually someone from the immediate environment, is hurt. All this is accompanied by a lack of sympathy with other people's suffering and a radical absence of fear. It is highly likely that these two attributes are the most important (57), firstly, because the memory traces generated by empathic reactions together with the outcomes of interactions with others, enable moral emotions of greater depth and scope to be crystallized than is the case through more fleeting connections. Moreover, the lack of fear prevents such people from developing a sense of anticipatory caution with regard to social punishment. As individuals therefore, they are decidedly unable to make powerful commitments (loyalty) and are not bothered by their own lack of compliance or transgression. There is thus a high probability of them engaging in predatory or shameless behavior, hence the concept "moral deficiency": such people engage in behavior that entails not only great risk to others but also, in the event that they are caught themselves — although some are capable of combining these unique emotional profiles with a gift for persuasion and an ability to lie that is so astute as to assure that they avoid being reprimanded in any way and end up holding influential positions in society.

It should be highlighted, however, that the moral deficits of desertion are strictly unilateral. They only share with psychopaths one side of the aspect of predatory behavior, one whose limits are marked out by doctrine. The numbing or annulment of empathic responses is applied only to those on the other side of the gap which separates them from the world of virtue, that is, members of the exo-group. Inside the realm of doctrine, however, their behavior is the exact opposite and becomes exemplary. Indeed, their fellow believers

expect to witness a commitment to the most tenacious and costly acts, which may culminate in self-sacrifice. Hence, those drawn to martyrdom are readily exploited by those who preserve, under all circumstances, selfish interests. As a rule, psychopaths do not usually squander their personal choices on grand commitments; on the contrary, their gift for simulating altruism is used for individual gain without any other kind of consideration. Potential martyrs, on the other hand, are by definition squanderers. Although they share with psychopaths some aspects of moral numbing the two phenomena are in fact quite distinct (29, 58, 59, 61). Whereas psychopaths are dangerous predators or pure villains by nature, martyrs tend to be gullible souls with a “beneficent” predisposition that nevertheless fails to prevent — indeed, in such cases it does just the opposite — them from leaving a host of victims in their wake upon offering themselves up as weapons of sectarian combat. Thus, what we are dealing with here is a form of deadly altruism, a collateral and highly damaging drift away from the cooperative tendencies of humans, and one that requires a detailed understanding.

Work Suggestions

We have characterized the phenomenon of the deadly martyr as a peculiar aspect of moral behavior. If this is indeed a suitable approach then the detailed description of brain mechanisms associated with morality may prove crucial. We have seen that there are areas and systems in the human brain that are responsible for processing moral sensibility, and also that these neural areas present a progressively demarcated functional specificity. Therefore, Darwin's criterion for the development of a moral sense and conscience in animals with well-founded social instincts is met, namely, a brain that is sufficiently developed to enable them to recognize their fellows, read their emotions, and remember and weigh up the outcomes of interactions. Consequently, the first step has already been taken. We now have available certain procedures that offer a tentative way in to that remote area of conscious self-scrutiny concerned with scruples and moral obligations. The complex road ahead involves teasing apart the subtleties of moral sensibility and its connections with neural mechanisms; it will be particularly important here to monitor the stages of moral sensibility during human development, and to study patterns of individual variability.

It is also likely that research will be carried out on evaluating moral discernment and its concomitant brain functioning in people selected for their heightened doctrinaire leanings (for traits such as loyalty, credulity, and radicalism). By this I mean that an experimental program must be developed

address the paradoxical co-existence of benumbing and moral exaltation that characteristic of those prepared to sacrifice themselves for the good of the group. Data is required, for example, on empathic perception with regard to suffering of members of the exo-group, comparing normative with sectarian individuals (50, 59). Likewise, it is imperative that we further our understanding of the links between biological markers of group identity and certain mechanisms of neurocognitive facilitation, as this interdependence is usually taken advantage of by doctrinal creeds in order to widen the gap between group and another (60, 63). This kind of evidence should serve to support, more strongly, the approach suggested in the present paper. Suffice to say this is a real possibility. In sum, I believe that it is worth taking seriously contributions of neuroscience in a field traditionally reserved for the social sciences, because our lack of knowledge about the forces which mold demagogues and martyrs and those who encourage them (4, 31, 35, 59) is proving far too costly.

References

1. Adolphs, R. (2003) Cognitive neuroscience of human social behavior, *Nature Reviews Neuroscience*, 4, 165-77.
2. Alexander, R.D. (1987) *The Biology of Moral Systems*, New York: Ald de Gruyter.
3. Anderson, S.W., Bechara, A., Damasio, H., Tranel, D. and Damasio, (1999) Impairment of social and moral behavior related to early damage human prefrontal cortex, *Nature Neuroscience*, 2 (11), 1032-37.
4. Atran, S. (2003) Genesis of suicide terrorism, *Science*, 299, 1534-9.
5. Blair, R.J.R. (1997) Moral reasoning and the child with psychopathic tendencies, *Personality and Individual Differences*, 22 (5), 731-39.
6. Bouchard, T.J., Lykken, D.T., McGue, M., Segal, N. and Tellegen, (1990) Sources of human psychological differences: the Minnesota study twins reared apart, *Science*, 250, 223-250.
7. Boyd, R. and Richerson, P.J. (1992) Punishment allows the evolution of cooperation (or anything else) in sizable groups, *Ethology and Sociobiology*, 171-95.
8. Boyd, R. and Silk, J.B. (2000) *How Humans Evolved*, New York: Norton.
9. Brosnan, S.F. and De Waal, F.B. (2003) Monkeys reject unequal pay, *Nature*, 425, 297-99.
10. Camerer, C.F. (2003) Strategizing in the brain, *Science*, 300, 1673-75.
11. Clutton-Brock, T. (2002) Breeding together: kin selection and mutualism in cooperative vertebrates, *Science*, 296, 69-72.
12. Davis, M.H., Luce, C. and Kraus, S.J. (1994) The heritability of characteristics associated with dispositional empathy, *Journal of Personality*, 62, 369-91.

13. Darwin, C. (1871) *The Descent of Man and Selection in Relation to Sex*, Princeton: Princeton Univ. Press (ed. 1981)
14. Dawkins, R. (1976) *The Selfish Gene*, New York: Oxford University Press
15. De Waal, F.B. (1996) *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*, Cambridge (Mass): Harvard University Press.
16. Eisenberger, N.I., Lieberman, M.D. and Williams, K.D. (2003) Does rejection hurt? An fMRI study of social exclusion, *Science*, 302, 290-2.
17. Fehr, E. and Gächter, S. (2002) Altruistic punishment in humans, *Nature* 415, 137-40.
18. Fehr, E. and Rockenbach, B. (2003) Detrimental effects of sanctions on human altruism, *Nature*, 422, 137-40.
19. Fehr, E. and Fischbacher, U. (2003) The nature of human altruism, *Nature* 425, 785-91.
20. Glimcher, P.W. (2003) *Decisions, Uncertainty and the Brain: The Science of Neuroeconomics*, Cambridge (Mass): The MIT Press.
21. Greene, J.D., Sommerville, R.B., Nystrom, L.E., Darley, J.M. and Cohen, J.D. (2001) An fMRI investigation of emotional engagement in moral judgment, *Science*, 293, 2105-8.
22. Goodall, J. (1986) *The Chimpanzees of Gombe: Patterns of Behavior*, Cambridge (Mass.): Harvard University Press.
23. Haidt, J. (2002) The Moral Emotions, in R.J. Davidson, K.R. Scherer and H.H. Goldsmith (eds) *Handbook of Affective Sciences*, New York: Oxford University Press (pp. 852-70)
24. Hare, R.D. (1999) *Without Conscience: The Disturbing World of Psychopaths Among Us*, London: Warner Books.
25. Heinsohn, R. and Parker, C. (1995) Complex cooperative strategies in group-territorial African lions, *Science*, 269, 1260-62.

26. Israeli, R. (2003) *Islamikaze: Manifestations of Islamic Martirology*, 1 Frak Croos and Co.
27. Kiehl, K.A., Smith, A.M., Hare, R.D., Mendrek, A., Forster, B.B., Brink and Liddle, P.F. (2001) Limbic abnormalities in affective processing by crim psychopaths as revealed by functional magnetic resonance imaging, *Biolog Psychiatry*, 50, 677-84.
28. Krebs, D.L. (2000) The Evolution of Moral Dispositions in the Hur Species, in D. LeCroy and P. Moller *Evolutionary Perspectives on Hun Reproductive Behavior*, Annals of the New York Academy of Sciences, 1 907, 132-48
29. Krueger, R.F., Hicks, B.M., and McGue, M. (2001) Altruism and antiso behavior: independent tendencies, unique personality correlates, dist etiologies, *Psychological Science*, 12, 397-402.
30. McCourt, K., Bouchard, T.J., Lykken, D.T., Tellegen, N. and Keyes, (1999) Authoritarianism revisited: genetic and environmental influer examined in twins reared apart and together, *Personality and Individ Differences*, 27, 985-1014.
31. Merari, A. (2002) *Terrorismo suicida* (Suicide Terrorism), Proceedings the conference: *Violencia, Mente y Cerebro* (Violence, Mind and Bra Valencia: Reina Sofía Centre for the Study of Violence, 157-169.
32. Millinski, M., Semmann, D. and Krambeck, H.J. (2002) Reputation h solve the “tragedy of the commons”, *Nature*, 415, 424-6.
33. Moll, J., De Oliveira-Souza, R., Eslinger, P.J., Bramati, I.E., Mou Miranda, J., Andreioulo, P.A. and Pessoa, L. (2002) The neural correlates moral sensitivity: a functional magnetic resonance imaging investigation basic and moral emotions, *Journal of Neuroscience*, 22, 2730-36.
34. Navarro, A. (2004) *Conflict and Cooperation in Human Affa* Dialogues: Proceedings of the Universal Forum of Cultures — Barcelona 2004.
35. Nisbett, R.E. and Cohen, D. (1996) *Culture of Honor: The Psychology Violence in the South*, New York: Harper and Collins.

36. Nowak, M.A. and Sigmund, K. (1998) Evolution of indirect reciprocity image scoring, *Nature*, 393, 573-77.
37. Phillips, M.L., Young, A.W., Senior, C., Brammer, M., Andrew, C., Calvert, A.J., Bullmore, E.T., Perret, D.I., Rowland, D., Williams, S.C.R., Gray, J.A. Davids, A.S. (1997) A specific neural substrate for perceiving facial expressions of disgust, *Nature*, 389, 495-8.
38. Preston, S.D. and De Waal, F.B.M. (2002) Empathy: its ultimate proximate bases, *Behavioral and Brain Sciences*, 25, 1, 1-72.
39. Quartz, S.R. and Sejnowski, T.J. (2002) *Liars, Lovers and Heroes: What the New Brain Science Reveals about How We Become Who We Are*, 1 William Morrow-Harper Collins Pub.
40. Queller, D.C., Ponte, E., Bozzaro, S. and Strassmann, J.E. (2003) Single gene greenbeard effects in the social amoeba *Dictyostelium discoideum*, *Science*, 299, 105-6.
41. Raine, A., Leucz, T., Bihrl, S., Lacasse, L. and Polletti, P. (2000) Reduced prefrontal gray matter volume and reduced autonomic activity in antisocial personality disorders, *Archives of General Psychiatry*, 57, 745-51.
42. Rands, S.A., Cowlshaw, G., Pettifor, R.A., Rowcliffe, J.M. and Johnstone, R.A. (2003) Spontaneous emergence of leaders and followers in foraging primate groups, *Nature*, 423, 34-37.
43. Ridley, M. (1996) *The Origins of Virtue*, London: Viking.
44. Rilling, J.K., Gutman, D.A., Zeh, T.R., Pagnoni, G., Berns, G.S. and Knetsch, C.D. (2002) A neural basis for social cooperation, *Neuron*, 35, 395-405.
45. Riolo, R.L., Cohen, M.D. and Axelrod, R. (2001) Evolution of cooperation without reciprocity, *Nature*, 414, 441-3.
46. Rushton, J. P. (1989) Genetic similarity, human altruism and group selection, *Behavioral and Brain Sciences*, 12, 503-59.

47. Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E. and Cohen, J. (2003) The neural basis of economic decision-making in the Ultimatum game, *Science*, 300, 1755-8.
48. Shaw, R.P. and Wong, Y. (1989) *Genetic Seeds of Warfare: Evolution of Nationalism and Patriotism*, Boston: Unwin Hyman.
49. Simon, H. (1990) A mechanism for social selection and successful altruism, *Science*, 250, 1665-8.
50. Sinervo, B. and Clobert, J. (2003) Morphs, dispersal behavior, genetic similarity, and the evolution of cooperation, *Science*, 300, 1949-51.
51. Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R.J. and Frith, C.D. (2004) Empathy for pain involves the affective but not the sensory components of pain, *Science*, 303, 1157-62.
52. Sober, E. and Wilson, D.S. (1998) *Unto Others: The Evolutionary Psychology of Unselfish Behavior*, Cambridge (Massachusetts): Harvard University Press.
53. Stephens, D.W., McLinn, C.M. and Stevens, J.R. (2002) Discounting reciprocity in an iterated Prisoner's dilemma, *Science*, 298, 2216-18.
54. Stern, P.C. (1997) Why Do People Sacrifice for their Nations?, in Comaroff and P.C. Stern (Eds.) *Perspectives on Nationalism and War*, Amsterdam: Gordon and Breach.
55. Stevenson, J. (2003) The Bin Laden effect: the new face of old-school terrorism, *International Herald Tribune*, December 3, (p.10).
56. Stone, V.E., Cosmides, L., Tooby, J., Kroll, N. and Knight, R.T. (2000) Selective impairment of reasoning about social exchange in a patient with bilateral limbic system damage, *PNAS*, 99, 17, 11531-36.
57. Tobeña, A. (2001) *Anatomía de la agresividad humana* (Anatomy of Human Aggressiveness), Barcelona: Galaxia Gutenberg.
58. Tobeña, A. (2004) Individual factors in suicidal attacks, *Science*, 304, 564-7.

59. Tobeña, A. (2004) *Mártires mortíferos: biología del altruismo letal* (Dea Martyrs: The Biology of Lethal Altruism), Barcelona: Tusquets Ed. (in press)
60. Tobeña, A., Marks, I.M. and Dar, R. (1999) Advantages of bias and prejudice: an exploration of their neurocognitive templates, *Neuroscience & Biobehavioral Reviews*, 23, 1047-58.
61. Tudge, C. (2002) Natural born killers, *New Scientist*, 2342, 11 May, 36-39
62. West, S.A., Pen, I. and Griffin, A.S. (2002) Cooperation and competition between relatives, *Science*, 296, 72-75.
63. Wilson, D.S. (2002) *Darwin's Cathedral: Evolution, Religion and the Nature of Society*, Chicago: Chicago University Press.
64. Wrangham, R. and Peterson, D. (1997) *Demonic Males: Apes and the Origins of Human Violence*, New York: Houghton Mifflin.
65. Wright, R. (1994) *The Moral Animal*, New York: Pantheon Books.