Values, Empathy, and Fairness across Social Barriers

EDITORS

Scott **ATRAN**

Arcadi **NAVARRO**

Kevin **OCHSNER**

Adolf **TOBEÑA**

Oscar **VILARROYA**

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES **VOLUME 1167**

ANNALS OF TH[illegible]DEMY OF SCIENCES
Volume 1167

Values, Empathy, and Fairness across Social Barriers

Edited by
Scott Atran, Arcadi Navarro, Kevin Ochsner, Adolf Tobeña, and Oscar Vilarroya

Published by Blackwell Publishing on behalf of the New York Academy of Sciences
Boston, Massachusetts
2009

The *Annals of the New York Academy of Sciences* (ISSN: 0077-8923 [print]; ISSN: 1749-6632 [online]) is published 32 times a year on behalf of the New York Academy of Sciences by Wiley Subscription Services, Inc., a Wiley Company, 111 River Street, Hoboken, NJ 07030-5774.

MAILING: The *Annals* is mailed standard rate. POSTMASTER: Send all address changes to *ANNALS OF THE NEW YORK ACADEMY OF SCIENCES*, Journal Customer Services, John Wiley & Sons Inc., 350 Main Street, Malden, MA 02148-5020.

Journal Customer Services: For ordering information, claims, and any inquiry concerning your subscription, please go to interscience.wiley.com/support or contact your nearest office:

Americas: Email: cs-journals@wiley.com; Tel: +1 781 388 8598 or 1 800 835 6770 (Toll free in the USA & Canada).
Europe, Middle East and Asia: Email: cs-journals@wiley.com; Tel: +44 (0) 1865 778315
Asia Pacific: Email: cs-journals@wiley.com; Tel: +65 6511 8000
Information for Subscribers: The *Annals* is published in 32 issues per year. Subscription prices for 2009 are:
Print & Online: US$4862 (US), US$5296 (Rest of World), €3432 (Europe), £2702 (UK). Prices are exclusive of tax. Australian GST, Canadian GST and European VAT will be applied at the appropriate rates. For more information on current tax rates, please go to www3.interscience.wiley.com/aboutus/journal_ordering_and_payment.html#Tax. The price includes online access to the current and all online back files to January 1, 1997, where available. For other pricing options, including access information and terms and conditions, please visit www.interscience.wiley.com/journal-info.

Delivery Terms and Legal Title: Prices include delivery of print publications to the recipient's address. Delivery terms are Delivered Duty Unpaid (DDU); the recipient is responsible for paying any import duty or taxes. Legal title passes to the customer on despatch by our distributors.

Membership information: Members may order copies of *Annals* volumes directly from the Academy by visiting www.nyas.org/annals, emailing membership@nyas.org, faxing +1 212 298 3650, or calling 1 800 843 6927 (toll free in the USA), or +1 212 298 8640. For more information on becoming a member of the New York Academy of Sciences, please visit www.nyas.org/membership. Claims and inquiries on member orders should be directed to the Academy at email: membership@nyas.org or Tel: 1 800 843 6927 (toll free in the USA) or +1 212 298 8640.

Printed in the USA.

The *Annals* is available to subscribers online at Wiley InterScience and the New York Academy of Sciences' Web site. Visit www.interscience.wiley.com to search the articles and register for table of contents e-mail alerts.

ISSN: 0077-8923 (print); 1749-6632 (online)
ISBN-10: 1-57331-760-8; ISBN-13: 978-1-57331-760-3

A catalogue record for this title is available from the British Library.

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

Volume 1167

Values, Empathy, and Fairness across Social Barriers

Editors

SCOTT ATRAN, ARCADI NAVARRO, KEVIN OCHSNER,
ADOLF TOBEÑA, AND OSCAR VILARROYA

This volume is the result of the first Barcelona Social Brain Conference, entitled **Values, Empathy, and Fairness Across Social Barriers: A Neurocognitive Approach to Fairness**, organized by the New York Academy of Sciences, the European Science Foundation, the Catalan Research Foundation, and the Social Brain Chair (Autonomous University of Barcelona), and held on November 21–22, 2008 at Fundació "la Caixa," Barcelona, Spain.

CONTENTS

Foreword

The moral virtues of humans are practised almost exclusively in relation to the men of the same tribe, and their opposites are not regarded as crimes in relation to the men of other tribes.

CHARLES DARWIN, 1882

This *Annals* volume collects most of the papers presented and discussed at the first Barcelona Social Brain Conference, entitled Values and Empathy across Social Barriers: A Neurocognitive Approach to Fairness, which took place on November 21–22, 2008. The conference was organized under the auspices of the New York Academy of Sciences (NYAS), the European Science Foundation, and the Catalan Research Foundation, as well as the Social Brain Chair (Autonomous University of Barcelona). The idea for the conference originally came about four years earlier, at the closing of an extremely large-scale set of multidisciplinary cultural debates, fostered by Dr. Joan Clos, who was then the mayor of Barcelona (and who is currently serving the Spanish ambassador in Turkey). It was Dr. Clos who suggested and lobbied for the creation of the Social Brain Chair, an institution devoted to disseminating current advances in the social neurosciences through education and research connected to "ruling human communities according to high standards in values and morality." In June 2006, the City of Barcelona and Autonomous University of Barcelona (UAB) convened to create the Chair, which was adscribed to UAB's Department of Psychiatry.

It has been gratifying to see the impressive media impact that the Social Brain Chair has had in promoting education on social-neuroscience issues in the metropolitan area of Barcelona, as well as the whole of Spain (see www.elcervellsocial.net). The Chair's initiatives include high-school contests concerning primatology, the production of a video documentary (*Bajo la Piel del Conflicto*, 2004), and academic publications such as *Social Brain Matters: Stances on the Neurobiology of Social Cognition* (O. Vilarroya and F. Forn-Argimón (Eds), Amsterdam: Rodopi, 2007).

Around this time, the Chair also contacted NYAS President Ellis Rubinstein, with a proposal to hold a series of conferences on social-brain issues, either in Barcelona or in New York City. The New York Academy of Sciences welcomed the initiative with enthusiasm. The present volume is a testament to two windy and hectic days in November 2008, on one of the steep slopes of Tibidabo mountain, where the Barcelona Museum of Science-"Cosmocaixa" is located. The conference venue came alive with a truly multidisciplinary atmosphere, where neuroscientists, economists, geneticists, anthropologists, and political scientists from around the world engaged in a highly productive encounter concerning the social neurobiology of fairness.

Values, Empathy, and Fairness across Social Barriers: Ann. N.Y. Acad. Sci. 1167: 1–4 (2009).
doi: 10.1111/j.1749-6632.2009.04879.x

The main topics of the conference (values, fairness, and empathy across social frontiers) were first outlined during a pleasant outdoor lunch at Scott Atran's summer cottage on the French Mediterranean coast, where Adolf Tobeña and Oscar Vilarroya had traveled in August 2007 to offer Atran the opportunity to lead the first Barcelona Social Brain Conference. While discussing findings from Atran and his collaborators on sacred values and their role in the Palestinian–Israeli conflict or the so-called, by then, "Global War on Terror" against Al-Qaida and similar organizations, it became clear that the notion of "social barrier" would be a good attractor for work coming from anthropological field research, neuroeconomics, and neurogenetic studies, as well as imaging experiments in social and affective neuroscience labs. With these ideas in mind, we contacted Kevin Ochsner in New York City, and he eagerly received the suggestion to lead one of the modules of the conference devoted to studies on empathy, with an emphasis on social distances, prejudices, and stereotypes. In Barcelona, Arcadi Navarro quickly committed to extend one the workshops he regularly organizes in order to head the "neuroeconomics" branch of the conference with a selection of game-theorists, geneticists, and evolutionary biologists. Furthermore, Scott Atran agreed to lead the section devoted to anthropological and political psychology studies. The final outline of the conference, including topics, speakers, schedules, and other details were decided at NYAS headquarters, during January 2008.

These distinctive branches, or modules, form a triad of pillars in a volume that we hope reads as engagingly as the talks presented at the conference on which each contribution is based. The introductory paper, by Adolf Tobeña, reproduces his keynote lecture that interconnects the conference topics and serves as a prelude to the present tome. Part I (empathy studies) was coordinated by Kevin Ochsner. It contains papers describing topics ranging from neural networks associated with mentalizing to characterizing brain regions mediating modalities of stereotyping. Part II (fairness studies) was coordinated by Arcadi Navarro. The papers in this section discuss genetic markers for fair/unfair proclivities in economic games, as well as the role of specific neuromodulators on social affects. Part III (sacred values studies) was coordinated by Scott Atran and contains papers devoted to the study of the role of sacred values in relation to maximal sacrifices during communal conflicts and the evolutionary roots of tribal instincts. The papers delivered by the different panels are complemented with a fine selection stemming from the posters presented at the conference.

The volume shows that the conjunction of rather disparate research fields is not only possible, but necessary for fine-grained neurocognitive depictions of the complex abilities sustaining the emergence and maintenance of human virtues (and sins). Merging data and models might improve the still developing power of multidisciplinary approaches to very nuanced phenomena. Whether this knowledge will serve one day to fulfill the aspirations of Dr. Joan Clos to foster "ruling human communities according to high standards of values and morality" remains to be seen. Nevertheless, we are confident that the impact of these seminal explorations along the frontiers covered by the book will be long lasting. In the final discussion that closed the conference, there were several participants who congratulated the sponsoring institutions for permitting such a varied and rich perspective, asking for a regular continuation of the initiative.

It must be said that the conference was also an opportunity for social gatherings. There were lively discussions after most presentations, and the attendants seemed to enjoy the socializing. The discussions continued and expanded while attendees enjoyed

the splendid and imaginative tapas food and drinks that galvanized participants (some 280 researchers from 24 countries!) during two stimulating and enriching days.

All this was the result of hard work and enthusiasm of many people who deserve praise. To be fair, we should acknowledge here all the people who have contributed in some way or other to make the Social Brain conference possible. However, this would take more space than the volume itself, so we will just mention those who have been directly concerned with the organization of the event. First and foremost, we would like to acknowledge the substantial institutional and personal support we received from the City of Barcelona, and especially from former Mayor Joan Clos, without whom neither the Social Brain Chair nor the conference would have come into being; the City Counselor, Montserrat Ballarín, for her unconditional backing of the Chair over the past two years; and former City Counselor Ramon Nicolau.

Second, we would like to recognize the support of the various institutions involved in the conference, especially the New York Academy of Sciences, the Fundació Catalana per a la Recerca i la Innovació (FCRi), the European Science Foundation, the Cosmocaixa and the "Obra Social" Fundació la Caixa, as well as the Autonomous University of Barcelona (UAB).

Third, we would like to thank the staff of the New York Academy of Sciences, who were committed to the organization of the conference, especially Kathy Granger, for her efficiency in the coordination of the activities at the NYAS site; Stacie Bloom, for her ideas and support through the entire process; and, of course, Ellis Rubinstein, for backing our project from the very first day we approached him.

Fourth, we want to thank the people at the FCRi, above all Jordi Mas, its executive director. He was committed to the conference from the very beginning; he was pivotal in finding the necessary institutional support for our endevour, reaching out to scientists, politicians, and business people alike, and making the overseas visitors feel at home. We also must acknowledge FCRi director, Joan Comella, who agreed to support the conference from the very first contact we had with him.

Fifth, we would like to recognize the collaboration of the Fundació "La Caixa" and Cosmocaixa personnel in the preparation and implementation of the conference, especially Beatrice Sala, Guillermo Santamaría, Anna Guil, and Jorge Wagensberg.

Last but not least, we would like to thank the Social Brain Chair team at UAB for making the conference possible. Josep Maria "Pepus" Rabanal conducted all the local and travel arrangements, and took charge of the onsite conference with efficiency, style, and without ever losing his smile. Ana Moreno coordinated all the conference programs and co-managed all the activities during the conference, doing it with ability and enthusiasm. Joseph Hilferty was unselfishly and totally commited in the preparation and implementation of the conference. We also thank Teresa Garzón, for preparation of the conference, and all the neuroimaging research group for helping in the registration processes, including Susanna Carmona, Erika Proal, Erika Barba, Elseline Hoekzema, and Olivia Merns.

Reference

1. Darwin, Charles. 1882. *The Descent of Men and Selection in Relation to Sex*, 2nd Edition. John Murray. London.

SCOTT ATRAN
University of Michigan
Ann Arbor, Michigan

ARCADI NAVARRO
Universitat Pompeu Fabra
Barcelona, Spain

KEVIN OCHSNER
Columbia University
New York, New York

ADOLF TOBEÑA
Universitat Autònoma de Barcelona
Barcelona, Spain

OSCAR VILARROYA
Universitat Autònoma de Barcelona
Barcelona, Spain

Lethal Altruists

Itineraries along the Dark Outskirts of Moralistic Prosociality

Adolf Tobeña

Autonomous University of Barcelona, Department of Psychiatry and Forensic Medicine, Bellaterra Barcelona, Spain

Suicide bombers are the most spectacular example of an impregnable morality toward one's own group that co-exists alongside a radical amorality toward members of another group. Suicide bombers carry out massacres with the utter conviction that they are acting in accordance with values associated with the greatest good. Suicidal attacks are conceived as a form of lethal altruism, a damaging drift from human cooperative tendencies and one that requires a detailed understanding. Strong altruism is a main component of a cluster of temperamental traits that may distinguish individuals with propensities to put themselves at the threshold of major progroupal sacrifices. Among all populations there will be pockets of extreme moralizing altruists willing to make high investments in others, investments involving great personal risk. A research framework is outlined to study other constitutionally based traits (dominance, boldness, aggressiveness, machiavellianism, narcissism, messianism, credulity/religiosity) that may also contribute to the different roles played by self-recruited members in combative cells that in turn are crucial for the ties they establish and the tactics employed. Individually oriented research may reveal profiles distinguishing between potential inducers and performers of martyrdom. As a rule, machiavellistic leaders do not usually squander their personal choices on group commitments; on the contrary, their gift for simulating altruism is used for individual gains. Potential martyrs, on the other hand, are by definition squanderers. Evidence accrued in recent years in fields going from behavioral economics to cognitive neuroimaging makes such an endeavor feasible.

Key words: **altruism; costly punishment; suicidal attacks; temperament; leadership; followership**

"Selfishness beats altruism within groups. Altruistic groups beat selfish groups. Everything else is commentary."[45]

On March 2003, Scott Atran published in *Science* an influential essay titled "Genesis of Suicide Terrorism".[1] At that time suicidal attacks had created deep alarm across Western countries to the point of dominating the strategic scenario. The strong apprehension only started to fade years later when it was clear that the launch of those onerous weapons by Islamic *guerrillas* would not shake the military options and status held by the big powers. Atran's approach was, by far, the best after the deluge of perplexed and mainly misguided interpretations that followed the September 11, 2001, attacks on the USA. On the grounds of sound data gathered in Palestine, Indonesia, Pakistan, and Sri Lanka, Atran claimed that suicidal attackers "had no appreciable psychopathology and were as educated and well-off as the surrounding populations." He attributed the main responsibility to "recruiting organizations that enlist prospective candidates from youthful populations" and characterized the roots of the phenomenon as "loyalty towards an intimate cohort of peers which charismatic trainers

Address for correspondence: Adolf Tobeña, Autonomous University of Barcelona, Department of Psychiatry and Forensic Medicine, School of Medicine Campus of Bellaterra, Bellaterra Barcelona 08193, Spain. Voice: +34935812783; fax: 34935811435. adolf.tobena@uab.es

Values, Empathy, and Fairness across Social Barriers: Ann. N.Y. Acad. Sci. 1167: 5–15 (2009).
doi: 10.1111/j.1749-6632.2009.04542.x

intensively cultivate within small cells," a deep psychological urge born from "sentiments of emotionally driven commitments that may have emerged under natural selection to override rational calculations during extreme challenges . . .sentiments which are purposely manipulated by leaders, recruiters or trainers to benefit the organization rather than the individual."

I wrote a letter to *Science* in response to Atran's essay with the aim of provoking a discussion because I had reservations concerning the purely contextual social-motivating approach favored by him. I asked friends for editing help to ensure that my response had an acceptable tone for North American sensitivity toward the topic, but no one replied for months. I had almost forgotten the issue when, on Christmas Eve 2003, I received an urgent message from an editor giving me 72 h to approve a version of the letter they had prepared (which I did as it only contained minor changes). That letter and Atran's reply were in print on April, 2, 2004.[2] The discussion had other avenues and participants and can be found at http://www.sciencemag.org/cgi/data/304/5667/47/DC1/1 to appreciate the degree of controversy.

Our main point of disagreement was Atran's rejection of individual traits as an explanatory source for suicidal attacks. He rejected that on the basis of the "fundamental attribution error"—the tendency to explain behavior in terms of individual personality traits when significant situational factors are at work. Such rejection, I argued, can lead to another fundamental miscalculation—that of neglecting temperamental traits laboring as propensities that make people more or less prone to devote themselves to maximal sacrifices. However obvious the institutional factors in any band that recruits and engages candidates to be used as weapons, they do not obviate the need to analyze whether the influence of peer pressure under a closed organization fully explains the exceptional behavior and motivation under scrutiny.

Emotionally driven loyalty in indoctrinated small cells had to be uncompromising during the preparation and execution of 9/11. Nonetheless, the sheer elaborateness of the plans make these attacks different from a low-cost bus ride by a single youth carrying a homemade bomb under his (or, less frequently, her) belt, with the aim of detonating the artifact in a crowded place. For the former you need trained and coordinated soldiers, for the latter any committed believer will do. This exemplifies that the paths leading to deadly martyrdom are diverse, although in my view they require not only extreme doctrinarian pressure and emotionally driven loyalties within cells cherishing "fictive kinships" but also differential talents and abilities that may help to distinguish between the roles played by individuals inside enclosed and combative cells.

Self-recruitment for high-risk experiences, for instance, is an individual attribute that often characterizes members of violent doctrinal groups, although it is not a distinctive feature among their members; fellow devotees, intermediate officers, and commanders typically share this attribute. But there are other constitutionally based traits (dominance, boldness, aggressiveness, machiavellianism, narcissism) that may make a contribution to the distinctive roles played by self-recruited members, which in turn are crucial for the ties they establish and the tactics they employ. In addition to the aforementioned traits, I proposed that costly punishment (strong altruism), messianism, and credulity/religiosity traits are individual propensities that might also contribute to the appearance of suicidal terrorism as a plausible option to harm the enemy in highly unequal contests.

Routes to Martyrdom: A Biological Scaffolding for Fanatic Drives

I elaborated on these ideas in a subsequent book, insisting that research directed to dissect activist's temperamental and cognitive profiles

may disclose hints to identify potential candidates for martyrdom as well as to distinguish between inducers and performers of suicidal attacks.[3] As a rule, machiavellistic leaders do not usually squander their personal choices on group commitments to the point of maximal sacrifice; on the contrary, their gift for simulating generosity and trustworthiness, through persuasive discourse, is used for individual gains without any other type of consideration. Potential martyrs, on the other hand, are by definition squanderers. Candidates drawn to martyrdom are readily exploited by those who preserve, under all circumstances, selfish interests. Despite sharing with their guides some aspects of moral benumbing, the two clusters are distinct. Whereas messianic sociopaths abound among commanders, martyrs tend to be gullible souls with a "beneficient" predisposition that nevertheless fails to prevent them from leaving a host of victims.

On the basis of the available knowledge, I advanced the following "working hypothesis" aimed at detecting main vectors behind suicidal attacks:[3]

- Strong altruism in the form of progroup loyalty (groupishness, parochial altruism) has biological roots and mediates costly helping behavior to the point of maximal sacrifice during communal contests. For proper functioning strong altruism needs reliable intragroupal recognition.
- There are primed signals of group identification (physical attributes, speech-intonation modes, ornaments, rituals) that induce biased neurocognitive processing. Intragroup signaling heralds the surge of parochialism while establishing preferential routes for indoctrination.
- Encapsulating beliefs (dogmatisms, sectarianisms, integrisms) become intertwined within these biased neurocognitive workings at the service of maximizing intergroup conflicts ("wars of ideas").
- There are differences in temperamental propensities toward parochialism, leadership, and agonistic friendship that explain the distinctive roles (leaders/followers; commanders/soldiers) that individuals spontaneously play in combative cells.
- There should be, as well, discernible neurocognitive profiles based on additional temperamental traits (machiavellianism, religiosity, self-confidence) distinguishing among doctrinarians—particularly between those who assume the task of fanaticizing from those others who host fanatic drives as devotees.

These are the steps or requirements upon which the narratives of highly cohesive and agonistic cells cultivate the *exceptional drive* leading to deadly sacrifice using members ("brothers," "comrades") as weapons. Several research endeavors have accrued fresh data that support most points of such theoretical scaffolding.

Costly and Parochial Altruists Everywhere

A variety of forms of strong altruism (punishment and investment in strangers at the cost of personal losses) have been intensively researched in the lab or in the field, establishing strong altruism as a trait characterizing a nontrivial proportion of human beings. Findings from different experimental settings devised by neuroeconomists have confirmed that there are pockets of individuals eager to: 1) punish defectors or violators of social norms at a personal cost; 2) reject unfair offers despite detrimental consequences for themselves; 3) punish unfair treatments when observing them as neutral and uninvolved third parties; 4) invest in strangers at a risk of losing goods.[4,5] Neuroimaging studies have identified several brain regions preferentially implicated in modulating these costly behaviors in games for real money—neural circuits that process fair versus unfair/unequal/disgusting options in social interactions.[6] Working in these neural networks can be altered as the result of neurological

lesions or by pharmacological interventions, such as oxytocin trust-enhancing effects.[7,8]

Strong altruism encompasses a wide variation of human cooperativeness escaping the restrictions of mechanisms based on helping kin (nepotism), the mutual interchange of goods (reciprocity), or the gains derived from building a reputation through image scoring (indirect reciprocity).[9] Strong altruism is easy to implement at the lab, using a variety of interactive games and has been found to operate on a wide range of human societies with different levels of technical and institutional development.[10] There are wide variations across cultures, but the results indicate that costly altruists, people prone to moralistic aggression and Samaritan generosity, never fail to appear.

The demonstration of an empirical association between strong altruism and parochialism (antipathy/hostility toward out-group members), represents a crucial junction that probably feeds the agonistic quality of loyalty between in-group members when conflicts between neighborhoods appear. To show automatic preferences for comrades, even transitory fellows in the well-known tradition of the "minimal group paradigm" of Tajfel,[11] is important but demonstrating a willingness to punish unfair strangers at a personal cost, particularly when the victims are in-group members, seems even more relevant for human naturalistic interactions. This is precisely what Ernst Fher'associates demonstrated working with 195 individuals (17–60-years old) from two Papua-New Guinea small-scale societies in the "third party punishing game." A neutral third party had to spend his money (KINA) to punish unfair "dictators" for violations of egalitarian sharing of 10 KINA gifts with unknown "recipients" in one-shot decisions.[12] Punishing appeared on every condition if dictators opted for not sharing at all or for blatantly ungenerous splits. Sanctions were, however, much higher when the three players were from the same group or when the recipient and third party were also comrades than when either the third party was an outsider (so viewing the unfairness from a detached frontier) or the recipient was the outsider (contemplating the unfairness applied to a foreigner).

In addition, in a long-range evolutionary simulation, Choi and Bowles found that altruistic and parochial (sectarian, ethnocentric, tribal) traits could have evolved jointly by promoting war between groups.[13] The simulation ran for 50,000 generations and allowed interactions between 20 groups ($n = 26$ members), with opportunities to trade, war, or no contact and fitness outcomes depending on the relative frequencies of combinations from two alleles (A/NA: altruists/selfish; P/T: parochial/tolerant) at two different loci, resulting in four types: PA (parochial altruists), TA (tolerant altruists), TNA (tolerant selfish), PNA (parochial selfish). The main pattern of results showed that there were only two stable populations dominated by parochial altruists (patriots thriving with frequent or at least occasional wars) and tolerant selfish (merchants devoted to profiting trade). Minor numbers of bullies (parochial selfish) or philanthropists (tolerant altruists) can survive under the warrior or trader's regimes but never predominate.

Parochialism is, of course, just a new word for the old notion of ethnocentrism that has been around for a long time, but these results are starting to illuminate the mechanism of selection that Darwin had postulated for the nonending competition and warfare between tribes or bands in human ancestry.[4,14]

Neurogenetics of Strong Altruism

After half a century dominated by a strong selfishness rule carrying almost the potency of a "diktat," geneticists and evolutionary biologists are cautiously redirecting efforts toward giving strong altruism an opportunity as motivator of human behavioral dispositions. The first attempts to measure the heritability of altruism as measured by self-report scales in twins were contradictory because genetic effects attained scores that varied from 50%[15] to none.[16]

Recent studies in large samples of twins using fair/unfair behavior in economic games have been much more consistent. More than 40% of rejection of unfair offers in the ultimatum game was explained by genetic effects in a study with 658 Swedish twins (comparing 71 dizygotic pairs with 253 monozygotic pairs).[17] In a subsequent work with these Swedish twins and USA sample recruited at a large twin gathering, investments in strangers during trials of the trust game provided heritabilities that ranged from 0.2 to 0.32, depending on the particular measure and subgroup.[18]

Regarding molecular targets and specific gene markers, an Israeli study showed an association between large allocations in the dictator game, with arginine vasopressin receptor 1a receptor gene long-repeat polymorphisms at promoter region RS3.[19] The same group had demonstrated, 2 years before, an association between dopamine receptor DRD4 variants with self-reported altruism.[20] Moreover, a British group in Cambridge demonstrated that disrupting 5-HT (serotonin) function through tryptophan depletion selectively alters reactions to unfairness without affecting mood, fairness judgement, reward processing, or response inhibition.[21] The transitory drop of serotonin was accompanied by an increase in rejection of unfair offers. A parallel finding was established by an Italian team measuring serotonin content in platelets;[22] normative volunteers with low serotonin scores had higher rejection of unfair offers. Both results cohere with many previous data showing subtle affiliative-agonistic modulation by serotonergic systems as well as with a study indicating that high basal levels of testosterone also enhances rejection and retaliation toward unfairness in the same game.[23]

It seems safe to conclude, then, that the trait of costly altruism and the behaviors that it encompasses (punish violators of social norms or proposers of unfair stakes, besides helping or investing in outsiders) are under a substantial genetic influence that may manifest itself through a variety of neurohormonal pathways and devices that have only started to be disclosed.

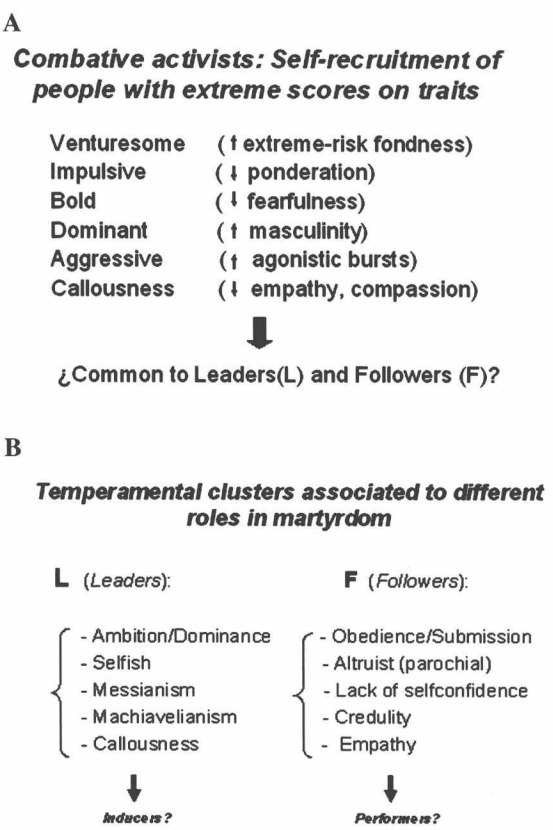

Figure 1. (**A**) Common temperamental traits for combative activists. (**B**) Potentially distinctive temperamental traits for leaders and followers within combative cells (from Ref. 2).

Temperament and Martyrdom: Vectors for Exceptional Characters

Despite the potential relevance of altruistic parochialism as a trait behind proneness to fighting and maybe martyrdom, it is necessary to combine this with approaches characterizing human temperamental profiles linked to prosocial/antisocial attitudes during communal conflicts.

Figure 1 depicts several temperamental traits fueling the attraction toward agonistic high-risk lifestyles—careers pursued, typically, by self-recruited volunteers entering violent gangs motivated either by territorial/profit making or by ideological (political/religious) goals. Bold, ambitious, dominant, and callous young males

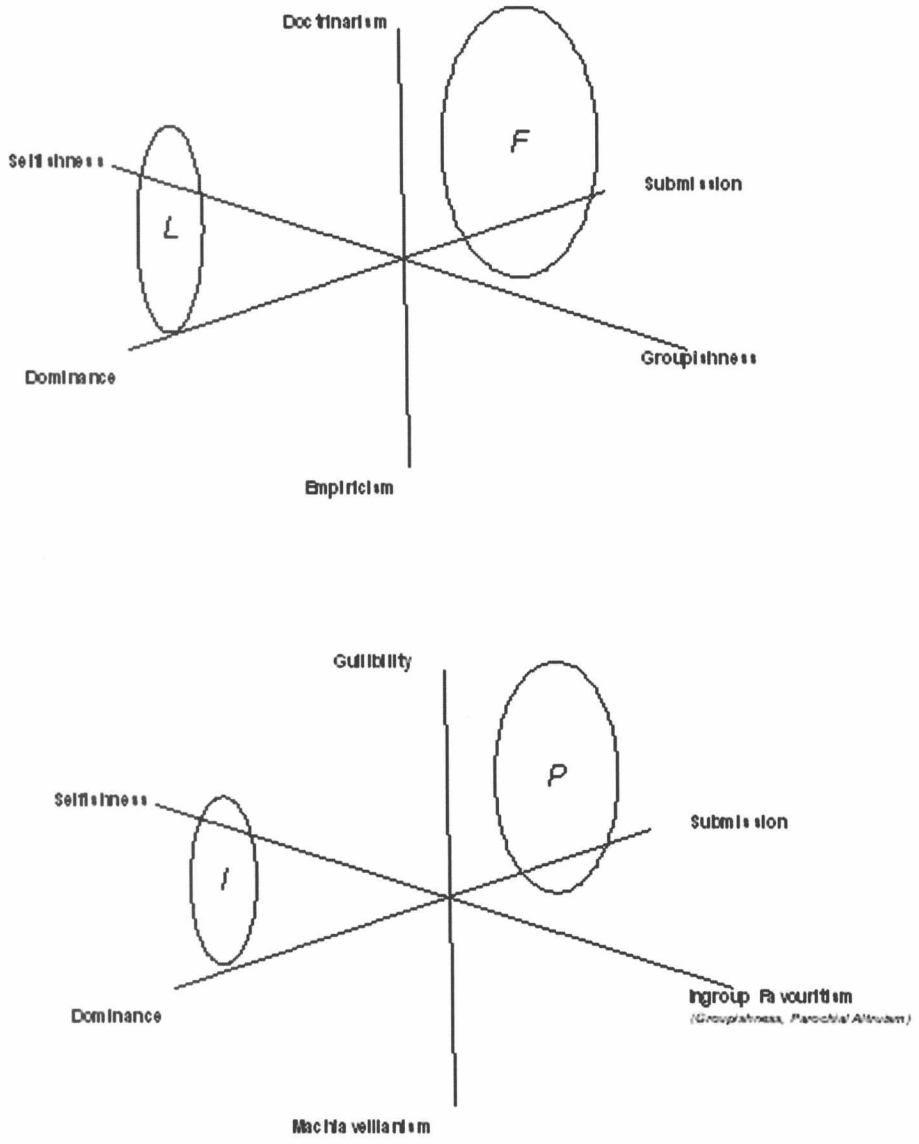

Figure 2. (**A**) Hypothetical space for agonistic clusters defining combative activists across three temperamental dimensions. (**B**) The empiricism/doctrinarism dimension in A is substituted by machiavellianism/gullibility). L, leaders/commanders; F, followers/soldiers. L and F clusters host inducers (I) and performers (P), respectively, of martyrdom actions (from Refs. 2, 25).

form a characteristic cluster in recruits of combative groups everywhere.[24,25] By adding the dimensions of selfishness/groupishness and machiavellianism/gullibility, potential distinctions between leaders/followers within combative bands can be deduced in the way suggested in Figure 1. Departing from these general traits, I also advanced a hypothetical dimensional space (Fig. 2) distinguishing between bubbles of potential

inducers and performers, depending on the weights of three dimensions: selfishness versus altruistic groupishness; dominance versus submission; empiricism versus credulity/doctrinarism.

I advanced these dimensions because there is a solid tradition of measuring dominance/ambition-leadership versus submission/conformity in personality research[26,27] as well as doctrinarism/empiricism either as such or in the form of machiavellianism/gullibility or tough/softmindedness.[28] Leadership heritability, for instance, has been calculated to be around 0.3 in studies with large samples of twins comparing life stories of leading in professional as well as leisure and friendship domains. Comparisons have also shown that leadership abilities depend more on social dominance and on ambition to achieve than on other temperamental characteristics.[29]

Such dimensional space is conjectural and needs to be assayed with the rest of traits as suggested. Studying the relevance of the empathy/callousness trait is compulsory in this regard, although the most needed are sensitive and consistent measures for selfishness versus groupishness because the existing scales for measuring altruism are a conflate of different tendencies.[16] Neuroeconomists have opened an alternative and very productive path: substitute (or better still, complement) psychometric scales with behavioral outcomes in interactive games for money[6] or status.[30]

These measures can be tentatively combined either with hormonal/physiological traits or with genetic markers as they have been used in pioneering studies of political preferences[31,32] as well as with neural signatures of temperamental styles studied through different neuroimaging/neurotracking methods.[30,33,34] Figure 3 depicts two proposals with plausible markers taken from studies that have brought molecular clues, as investigated in animal models of social behaviour, to the human social landscape.[35,36]

Neurocognition of In-Group Favoritism: An Agenda for Neuroimaging Research

The agonistic drift of altruistic parochialism can be fruitfully explored from another perspective by studying the subtleties of prejudices and stereotypes toward outsiders, as described by social psychology, with neuroimaging equipment. Susan Fiske's study at Princeton exemplifies this.[37] Fiske has been able to demonstrate that, when students view pictures of out-group individuals, the brain activation/deactivation patterns on functional MRI (fMRI) scans differ depending on the specific attributes of such outsiders. Pictures of extreme out-group persons situated on a low-warmth–low-competence quadrant (i.e., disgusting homeless and addicts) induce activations of the anterior insula and amygdala but not of the medial prefrontal cortex (mPFC), meaning that the threat–aversion brain systems are highly reactive but devoid of the signature of an interaction with another person, which is an mPFC-modulated function. Extreme outgroups are thus automatically dehumanized by normative people during social interaction routines. It would be extremely informative to explore if similar automatic dehumanizing occurs with individuals tagged as foreign enemies, comparing fMRI activations of parochial altruists versus selfish tolerants, for instance.

The array of behavioral games devised by neuroeconomists has been transferred to the neuroimaging lab and can also be used to study how and what neural systems process differential reactions to in-group/out-group signals. The well-characterized implication of anterior insula in rejection of unfair offers in the ultimatum game,[38] the ventral striatum modulation of gratifying vengeance obtained either through monetary sanctions to unfair players in the trust game,[39] or by inflicting pain (faked electric shocks) to unfair players after a public goods game[40] are good candidates. Neuroscans during realistic retaliation by administering mechanical pain to unfair opponents in an