Advances in Complex Systems, Vol. 14, No. 5 (2011) 711–731

DOI: 10.1142/S0219525911003190

AGENT-BASED SOCIAL PSYCHOLOGY: FROM NEUROCOGNITIVE PROCESSES TO SOCIAL DATA

NESTOR CATICHA*,‡ and RENATO VICENTE†,§

**Departmento de Fisica Geral, Instituto de Física, Universidade de São Paulo, CP 66318, 05315-970, São Paulo - SP, Brazil*

†Department of Applied Mathematics, Instituto de Matemática e Estatística, Universidade de São Paulo, 05508-090, São Paulo - SP, Brazil

‡nestor@if.usp.br

§rvicente@ime.usp.br

Received 29 September 2010
Revised 26 July 2011

Moral Foundation Theory states that groups of different observers may rely on partially dissimilar sets of moral foundations, thereby reaching different moral valuations. The use of functional imaging techniques has revealed a spectrum of cognitive styles with respect to the differential handling of novel or corroborating information that is correlated to political affiliation. Here we characterize the collective behavior of an agent-based model whose inter individual interactions due to information exchange in the form of opinions are in qualitative agreement with experimental neuroscience data. The main conclusion derived connects the existence of diversity in the cognitive strategies and statistics of the sets of moral foundations and suggests that this connection arises from interactions between agents. Thus a simple interacting agent model, whose interactions are in accord with empirical data on conformity and learning processes, presents statistical signatures consistent with moral judgment patterns of conservatives and liberals as obtained by survey studies of social psychology.

Keywords: Agent-based model; opinion dynamics; reinforcement learning; statistical mechanics; sociophysics; neurosociology.

1. Introduction

The proponents of Moral Foundation Theory (MFT) [36] have identified at least five moral foundations or dimensions that, potentially, are universally present in humans. These dimensions are manifested in different manners, not only across time and cultures, but also within a society. Individuals with different attributions of the relative importance of the dimensions will be led to fundamental misunderstanding of moral motivations of each other. Within the MFT, extensive empirical support [31, 35, 37] has been gathered for the fact that the use of different subsets of moral foundations by groups is significantly correlated with a scale characterizing the group along the political spectrum. The subsets are such that liberals tend to rely more strongly on aspects relating to (a) harm/care and (b) fairness/reciprocity. Conservatives rely on these aspects but not as much. In addition they also regard as important (c) in-group loyalty, (d) authority/respect and (e) purity/sanctity, to a larger extent than liberals [36].

At the scale of individuals, empirical evidence supports the following. (i) Moral opinions are to a large extent emitted automatically, that is, people, as a first approximation, are intuitionists [32–35]. (ii) There is a psychological cost of dissent, with humans trying to attain social conformity modulated by peer pressure [7, 21, 54, 56]. (iii) Conformity is learned from interactions within a social network [41]. (iv) Individual cognitive strategies may differ with respect to the relative sensitivity to learning from novel information as compared to reinforcing habitual responses [6].

Despite the growing body of experimental evidence accumulated over the last decade, explicit connections between this new empirical evidence on individual behavior and social phenomena (or between micromotives and macrobehavior [51]) still are relatively unexplored. Since the work on sociophysics by Galam [25, 26], the statistical mechanics community has already addressed this aggregation problem in social systems [14, 20, 28, 29, 50, 64].[a] However, this research has mainly focused on the study of simplified scenarios based on common sense suppositions that result in models that are interesting *per se*. Some recent works have exemplified a different direction by postulating reasonable inter-individual interactions and trying to predict [27] or explain [9] empirically observed aggregate behavior. We believe that the program of building models explicitly based on the empirical evidence that is now available [24] is worth pursuing.

We do not believe that any stylized model can provide an all encompassing and precisely quantitative description of human nature. Our general goals in this work are instead much more modest and can be stated as follows. Our first goal is to provide a mathematical model that is capable of connecting in the same framework empirical evidence from processes at different scales. We also would like to have a framework that is capable of instigating the formulation of new theoretical and experimental questions. In particular, we propose a model with agents consistent with empirical evidence and study its aggregate behavior by employing the approach and numerical techniques of statistical mechanics. We then show that this aggregate behavior predicts that a well-defined feature is expected to be observed in the data we are considering. We then verify the consistency of our predictions and propose a new interpretation to empirical evidence within this framework that can be qualitatively tested against new data sets in the future. We insist that such a general model is only expected to provide qualitative predictions of limited scope

[a]That methods from physics can be used in sociology is not particularly surprising, as these methods were actually brought to physics in the 19th century by James Clerk Maxwell when he was inspired by the historian Henry Buckle's account of Adolphe Quetelet's statistical approach [49] to social science ([12, p. 438]).

and emphasize our belief that a good model should make testable predictions but should not explain too much.

In the following sections, we give details on the empirical evidence we consider relevant, introduce our modeling approach, present and discuss the results obtained. A brief summary of methods employed is provided as an appendix. Computer programs used as well as the data set analyzed are also made available in [16].

2. Empirical Evidence for the Model

In building our model, we have tried to incorporate in a stylized manner empirical evidence. In this section, we describe what we believe to be essential empirical observations and which model structures they suggest.

2.1. *Moral theories and the automaticity of moral judgments*

Philosophers have struggled with the problem of conceptualizing morality since antiquity. Although philosophical theories tend to be of normative character they can be regarded as a starting point in considering a possible scientific approach for morality as a social phenomenon. Three theories are of particular interest to our discussion: virtue theory, deontology and utilitarianism [11, 38].

According to [13, 38], utilitarianism proposes that moral judgments should be based on the consequences resulting from them. Only actions that maximize social happiness and minimize pain should be taken. In the deontological view only actions that could be universally adopted without violating anyone's rights should be pursued. Virtue theory takes into account the intrinsic limitations of human nature and states that morality is concerned with maximizing virtues and minimizing vices. Each view of morality presupposes cognitive loads that can be experimentally verified. While utilitarianism and deontology concentrate moral decisions on higher cognitive functions in the prefrontal and sensorial brain regions, virtue theory proposes the coordinated functioning of these areas with others associated to the processing of emotions. Data gathered in the last decade favors a combination of the three views with preponderance of a mode that is closer to virtue theory [34, 42]. Actually ample research points in the direction that moral opinions are mostly formed with very few recourse to utilitarian (or consequentialist) reasoning.

Evidence supports that moral violations elicit strong negative responses that activate the socio-emotional structures of the brain (e.g. medial prefrontal cortex and posterior cingulate cortex).[b] But emotions are not the only component behind every moral judgment as experimental observation also suggests that these automatic negative responses can be overcome by a more utilitarian mode by recruiting cognitive areas in the prefrontal cortex, in particular, when difficult personal dilemmas with important social consequences are involved [33, 42, 48].

[b] For a discussion of the emotional components of moral intuitions and their neural substrates see [45, 65], see [34, 32] for fMRI evidence and [35] for psychological tests.

In this work, we simplify by only considering the moral grading of statements and not the comparison and subsequent choice between different possibilities and their consequences. Therefore we keep from entering a discussion between deontological or consequentialist [11] moral theories which might guide the modeling of decisions and choices associated to moral dilemmas. We start by supposing as a first approximation that socio-emotional intuitions predominate and that moral grading is in fact automatic.

2.2. *Moral foundations*

Human culture and values are markedly diverse. Nevertheless, this diversity seems to emerge from innate universals. Modern research in cultural anthropology [55], primatology [19] and evolutionary psychology [36, 40] suggests that morality may be parsed into a small number of basic intuitions. Haidt and collaborators [31, 36, 38, 37] reviewed the literature to identify five candidates to innate moral intuitions (or foundations) associated to: care, fairness (classified as individualizing foundations), loyalty, authority and purity (binding foundations). This set of innate moral foundations could have coevolved with culture due to adaptive challenges primate populations have been subjected to in their evolutionary history [10, 36]. This innateness, however, does not imply moral judgments that are rigid or genetically determined. What is considered to be a virtue or a vice in a given society at a given time depends instead on learning and imitation in a social environment. This plasticity from an initial draft is the key to understand how diversity can be universality-bound [31, 44]. In our model, we introduce moral foundations as dimensions in an abstract moral state space (for a similar suggestion, see [17]). Five-dimensional moral vectors live in this space and are animated by an adaptation dynamics elicited by social interactions.

2.3. *Reinforcement learning*

Extensive literature (see [39] and references therein) suggests the existence of a generic machinery for error and conflict processing in humans. This adaptive circuit implements a full-fledged reinforcement learning system. In this system, the basal ganglia processes error information provided by the spinal cord, the sensorial cortex and by areas that were traditionally labeled as the limbic system. This error measure is then converted into a dopamine signal that is used to correct responses with the mediation of a strategically connected region known as the Anterior Cingulate Cortex (ACC).

In event related potential (ERP) experiments, the processing of social exclusion feelings and social normative conflicts has been associated to error related negativity (ERN) signals with source located in the ACC. This localization has been further confirmed by fMRI [39]. Initial activation of the ACC due to conflict has also been associated to subsequent alignment to opinions perceived as preponderant in the social group [41]. These new data corroborate classical behavioral experiments on the psychology of conformity conducted by Sherif [54] and Asch [7]. Such findings suggest a central role to reinforcement learning in the dynamics of conformity to social norms. The observation of amigdala activation during social conflict [8] together with the known association of the ACC activation when physical pain is involved [21, 56] additionally suggest that disagreement elicits a psychological cost in humans.

Both the automaticity of moral judgments and the psychological cost of disagreement can be represented by using a reinforcement learning model that is well-established in computer science [58] and statistical mechanics [22]. Within this model a moral judgment is regarded as a classification task. Each agent has an internal moral state $\mathbf{J}_i$. At each timestep, an agent is chosen and its internal state is updated to minimize the psychological cost. If there is no noise in the communication, this minimization follows a gradient descent dynamics:

$$\begin{aligned} \tilde{\mathbf{J}}_i(t+1) &= \mathbf{J}_i(t) - \epsilon \nabla_{\mathbf{J}_i(t)} \mathcal{H}, \\ \mathbf{J}_i(t+1) &= \frac{\tilde{\mathbf{J}}_i(t+1)}{|\tilde{\mathbf{J}}_i(t+1)|}, \end{aligned} \tag{1}$$

where ϵ defines the time scale and $\mathcal{H}$ represents the social cost, namely, the sum of psychological costs incurred by an agent in a given social network.

2.4. *Cognitive styles*

A recent experiment [6] has shown evidence that there is a correlation between being a liberal or conservative with respect to social issues and the way novel or corroborating information is used. The experimental setup consists of the measurement of ERPs and concomitant fMRI while participants are exposed to a Go/No-Go task. The subjects first habituate to a frequent "Go" stimulus. In some rare occasions, a "No-Go" stimulus appears and a related ERN signal is registered. As expected, a localization algorithm and simultaneous fMRI identify the ACC as a source for the conflict signal. Before the experimental section, the participants are asked to rate their political orientation from -5 for "very liberals" to $+5$ for "very conservatives". A negative correlation is then found between political affiliation and the amplitude of the ERN signal. Liberals exhibit more intense conflict related activation of the ACC as compared to habituated response.

Two other recent studies provide further evidence by associating political behavior and genetic differences affecting dopamine receptors known as DRD2 and DRD4 [18, 53]. Dopamine is a neurotransmitter directly related, among several other things, to predictive reward systems that modulate reinforcement learning mechanisms [52].

We regard these experiments as suggesting that self-declared liberals are more at ease with novel information and rely less on corroborating information while self-declared conservatives prefer corroboration and are less at ease with novelty. We use the term corroboration to signify "confirming and in accordance to previous opinion". This new empirical work concurs with ample literature in social sciences which have discussed for decades the relation between cognitive styles and political orientation [2]. Whether this cognitive diversity is due to genetic or cultural conditions is beyond our present scope.

We try to capture some aspect of the information conveyed by this class of empirical results by recurring to models of statistical learning [22]. We propose that there are different learning styles according to the balance between novelty seeking and corroboration. For that we introduce a parameter $0 \leq \delta \leq 1$ that specifies the amplitude ratio between learning corroborating and learning conflicting information and define a learning algorithm interpolating between pure corroboration learning when $\delta = 1$ and pure novelty seeking learning when $\delta = 0$. In one social interaction, we would then have

$$\begin{aligned} \tilde{\mathbf{J}}_i(t+1) &= \mathbf{J}_i(t) + \epsilon F(h_i, h_j)\mathbf{x}, \\ \mathbf{J}_i(t+1) &= \frac{\tilde{\mathbf{J}}_i(t+1)}{|\tilde{\mathbf{J}}_i(t+1)|}, \end{aligned} \tag{2}$$

where the function

$$F(h_i, h_j) = \begin{cases} \delta h_j & \text{if } h_i h_j > 0 \\ h_j & \text{otherwise} \end{cases}, \tag{3}$$

modulates learning of agent i by comparing its classification $h_i = \mathbf{J}_i \cdot \mathbf{x}$ of an input $\mathbf{x}$ with that issued by a social neighbor h_j. The response of agent i is then corrected with direction given by the issue and sign and amplitude dictated by the information provided by agent j.

2.5. *Social influence*

Classical experimental set-ups by Sherif [54] and Asch [7] demonstrated that groups influence individual beliefs and decisions. In Sherif's experiment, subjects are placed in a dark room and asked to judge the displacement of a spot of light without knowing that it is actually stationary. The task is repeated a number of rounds with participants either alone or in groups. When in group, individual estimates converge to a group specific norm.

In Asch's experiment, a participant is placed in a room with a group of other people that are, without her knowledge, confederates of the study. She is then presented with two cards. One with a standard vertical line and the other with one vertical line the same length of the standard and two with different lengths. The participant is then asked to identify which of the lines in the second card is most similar to the standard line in the first card, but this is done after all confederates unanimously make the wrong choice. A strong conformist trend is observed, with decreased accuracy of individual judgment.

These experiments are generically consistent with the modern picture of conflict mediated reinforcement learning [41]. However, two aspects of these set-ups require a closer examination: in the first set-up participants are anonymous and information ambiguous and in the second information is comparatively objective and subjects largely uncategorized.

Modern social psychology defines three types of social influence [1]: informational, normative and referent informational. Informational influence is the predominant mechanism for social influence in the absence of objective evidence and when no group identification is present as in Sherif's experiment. The expectation of acceptance or punishment by other members in a group leads to normative influence observed in Asch's experiment. When group membership is salient referent informational influence becomes dominant. In this mode, individuals seek to be identified as pertaining to a given group.

To investigate referent informational influence in [1] Sherif's and Asch's experiments are repeated with the introduction of salient group membership. It is observed that if a participant regards herself as part of a different group, the conformity effect is greatly diminished.

In our modeling effort, we then suppose that referent informational influence is preponderant and, as a starting point, assume the extreme scenario where only in-group conflict leads to conformity effects. We therefore start by considering the case in which agents are circumscribed to social neighborhoods with homogeneous cognitive styles represented by the corroboration/novelty parameter δ.

2.6. *Social topology*

Social networks have received a great deal of attention during the last decade [5]. The Internet and social networking sites like Facebook, MySpace or Orkut now make possible to study empirically topological and dynamical properties of social graphs. One of the simplest topological properties that can be defined is the distribution of node degrees $P(k)$. A growing number of studies seems to indicate that many natural networks are well-represented by scale-free graphs with the tail of the node distribution given by $P(k) \sim k^{-\gamma}$ [5].

To make an informed modeling choice, we have searched the literature for networks representing social interactions. We have found some illustrative cases. For instance, the node distribution of a network of phone calls has been found to be scale-free with $\gamma = 2.1$ in [4]. The network of sexual contacts has been identified as being scale-free with $\gamma = 3.4$ in [43]. By employing a sampling algorithm, a power law with $\gamma = 3.4$ has been reported for the Facebook [30], but this estimate relies on a range of degrees spanning only one decade. Other study of entire networks instead supports an exponential node distribution in this case [59]. A third study [3] with samples from three social network websites reports scale-free behavior with $\gamma = 2$, $\gamma = 3.1$ and $\gamma = 3.7$, respectively for Cyworld, MySpace and Orkut.

We have considered empirical evidence and have applied the simple procedure of preferential attachment described by Barabási and Albert [5] to generate scale-free social networks with $\gamma = 3$. We are aware of the fact that clustering properties of networks built in such way will differ from those found in real social networks, however, we have been able to verify that the particular results we present here are sensitive to the degree distribution and qualitatively robust in relation to other topological properties [62].

3. Agent-Based Model

3.1. *Combining empirical ingredients*

We introduce a model for an interacting society where agents represent individuals that debate moral issues with their social neighbors. In general terms our modeling approach continues a now established line of research on opinion dynamics [14, 28, 61]. We however, strongly emphasize that information exchanges and processing, even though stylized, should be explicitly linked to the empirical evidence available.

We start by supposing that the moral state space has $M_D = 5$ dimensions so that moral issues may be parsed into these dimensions. We simplify by assuming only unit vectors. It is certainly possible that the same results we have reached could have been obtained by assuming less (or more) dimensions, however, it would be incompatible with known empirical data that supports the existence of five moral dimensions [35–37].

We consider that an agent i attributes a moral content for an issue μ that may then be represented by a five component unit vector (*issue vector*) [61] $\mathbf{x}_{i\mu} = \mathbf{x}_\mu + \mathbf{u}_{i\mu}$ with $\mu = 1, \ldots, P$ and $i = 1, \ldots, N$. Here $\mathbf{x}_\mu$ represents the average part of the moral parsing and $\mathbf{u}_{i\mu}$ represents an individual part.

We call the average (normalized) issue $\mathbf{Z} \propto \sum_{i=1}^{N} \sum_{\mu=1}^{P} \mathbf{x}_{i\mu}$ the *Zeitgeist vector*, which can be regarded as describing the cultural environment and providing a symmetry breaking direction in the moral state space. Here we are not to be concerned with the origin of the *Zeitgeist*[c] vector as it results from evolutionary and historical processes taking place in time scales that exceed the scope of our simple model. We further simplify the model by assuming that individual components are such that $\sum_{i=1}^{N} \sum_{\mu=1}^{P} \mathbf{u}_{i\mu} = 0$ and are small enough so that they can be disregarded in a first analysis. We are also going to assume that only the average part is correlated over the social network.

The relevant variables to characterize an agent are suggested by moral foundation theory. For each agent and unavailable to other agents, the internal moral state is encoded in another unit vector $\mathbf{J}_i$ (*moral vector*), also five-dimensional, the magnitude of each component representing the weight the ith agent gives to a particular moral foundation. Unit vectors are used to avoid introducing the collateral notion that one agent could be more moral than another.

The automaticity of moral judgment is then represented as a classification task where in an elementary interaction an agent i gathers information on the moral

[c]For an essay on the role of the *Zeitgeist* in historical explanation, see [23].

classification of her social neighbor j on a given issue μ. This classification is represented by a field given by $h_{j\mu} = \mathbf{J}_j \cdot \mathbf{x}_{j\mu}$. Its attributes are its sign and magnitude, indicating respectively whether the issue is considered morally acceptable ($h_{i\mu} > 0$) or not ($h_{i\mu} < 0$) and how strongly the agent holds this position ($|h_{i\mu}|$).

At this point, we make an additional abstraction leap that leads to a still simpler model: we suppose that a debate is a more complex interaction that involves multiple issues and multiple agents and it works effectively as if the participants were estimating the *Zeitgeist* vector $\mathbf{Z}$. Therefore, the effective interaction we are going to consider corresponds to the exchange of fields $h_j = \mathbf{J}_j \cdot \mathbf{Z}$ between neighbors.

We only consider here social influence between similar cognitive styles, namely, agents have homogeneous cognitive styles and interactions are symmetric. To consider the empirical fact that in-group disagreement (or being in the minority) elicits a negative brain response we introduce a measure of the psychological cost of disagreement between socially interacting agents i and j. It is quantified by $V_\delta(h_i, h_j)$, a function of their opinions which depends on a parameter δ that measures the different treatment of corroborating or novel opinions.

Reinforcement learning can be recast in its off-line version [22] as the process of seeking a minimum in a given cost landscape. Along this line we assume that moral vectors $\mathbf{J}_i$ evolve by decreasing the psychological cost under communication through a noisy channel. The social cost $\mathcal{H}$ is defined by summing V_δ over all pairs of interacting agents (i, j) that are edges of the social graph:

$$\mathcal{H}(\{\mathbf{J}_i\}) = \sum_{(i,j)} V_\delta(h_i, h_j). \tag{4}$$

It depends on $\{\mathbf{J}_i\}$, the configuration of the society, and on the cultural environment, given by the *Zeitgeist* vector.

The functional form of the psychological cost must reflect experimental data. The only stylized fact we include is that there are different cognitive styles regarding the different way that novelty and corroborating data is handled. By integrating the modulation function proposed in Eq. (3), we reach a reasonable choice, by no means unique, that is depicted in Fig. 1,

$$V_\delta(h_i, h_j) = \frac{1}{2}(1-\delta)|h_i h_j| - \frac{1}{2}(1+\delta)h_i h_j. \tag{5}$$

The corroboration/novelty parameter $\delta (0 \leq \delta \leq 1)$ quantifies a cognitive strategy with respect to the difference in treatment of agreement and disagreement as it is suggested by ERP experiments. Our agents are conformists, namely, in the face of disagreeing opinions, the dynamics is such that the social cost is decreased. For $\delta = 0$ (leftmost panel of Fig. 1) agents are novelty seekers and do not use corroborating opinions since $V_{\delta=0}(h_j, h_k)$ is flat for opinions of the same sign. For $\delta = 1$ (rightmost panel of Fig. 1), agents seek corroboration and conformity, learning equally in the case of agreement or disagreement.

The techniques of statistical mechanics permit obtaining collective or aggregate emergent phenomena arising from reinforcement learning with the value of the social

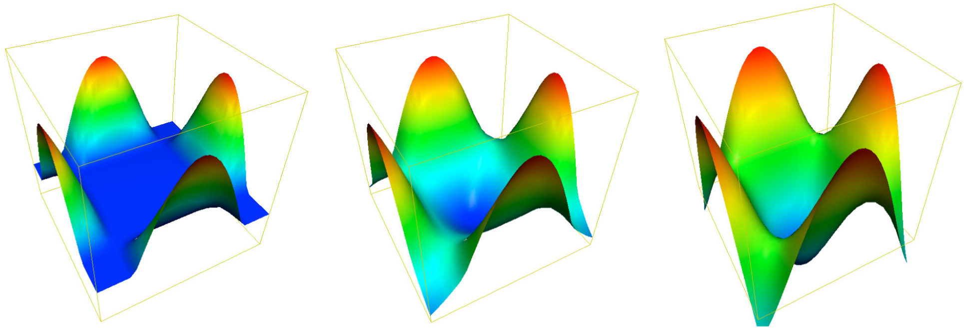

Fig. 1. Psychological cost: $V_\delta(h_j(\theta_j), h_k(\theta_k))$ as a function of θ_j and θ_k, the angles between $\mathbf{J}_j$, $\mathbf{J}_k$ and the *Zeitgeist* vector $\mathbf{Z}$ where $h_j = \cos\theta_j$. The potential can be written as $V_\delta(h_j, h_k) = -h_j h_k$ if $h_j h_k < 0$ (disagreement) and $V_\delta(h_j, h_k) = \delta h_j h_k$ if $h_j h_k > 0$ (agreement). The figure depicts cases with $\delta = 0$ (left), $\delta = 0.4$ (center) and $\delta = 1.0$ (right). The noisy learning dynamics tends to change the $\mathbf{J}$ making V_δ decrease along its gradient. Four peaks represent the cost of maximum disagreement when moral state vectors are opposite and angles are $(0, \pm\pi)$ and $(\pm\pi, 0)$. Note that when agents agree about the sign of their opinions, the benefit of agreement increases with δ.

cost $\mathcal{H}$ or at least, its average, constituting relevant information to characterize the state of the society with respect to the current *Zeitgeist*.

3.2. *Incomplete information, statistical mechanics and peer pressure*

The incompleteness of the available information about agent moral vectors imposes the use of probabilities in describing the moral state of the society. What is the information available to construct the theory? From the evidence about cognitive styles of persons we have constructed two functions. First, the psychological cost, Eq. (5), which describes the cost of disagreeing as a function of the opinions of the agents. Second, the social cost, Eq. (4), is the sum over pairs of interacting agents of the psychological cost. The social cost is about the whole society and results obtained from it will include collective properties arising from the interaction among the agents. It carries two types of information, about the internal space of pair interactions or cognitive level and about the external space, specifically about the geometry of the neighborhood of interactions, or the social level.

A complex system such as a society can be described in many ways. Suppose that we choose to study experimental questions which will have the same answer when the social cost has a given value or a given expected value, averaged over the probability distribution. There might be questions that do not fall into this category. In a physics language, an experimentalist wishes to prepare a system, by deciding on the control of certain parameters, in such a way that repetitions of the experiment will result in compatible answers for a class of questions. There will be other questions that, not resulting in predictable answers, cannot be addressed within that experimental set-up. This might be because, at last they are not interesting, or that another experimental design is needed in order to examine them. We concentrate on those questions for which knowing the expected value of the social cost is sufficient. If this cost is not known, then the most tempting thing to do is, by claiming insufficient reason to pick one direction over other, is to assign a uniform distribution for the set of moral vectors. Now, upon learning that the social cost of the society is an important quantity that defines the state, we suppose it is known. This simple assumption leads to the introduction of a conjugate variable, the peer pressure scale. Knowledge of one permits calculating the other, although this might be very difficult to do. At any rate, if we ignore both, for a given experimental system, the theory requires that one of them be measured. This is how it goes. We start from a uniform distribution $P_0(\{\mathbf{J}_i\})$. Suppose that new information is obtained, now the expected value of $\mathcal{H}$ is known. This is the average with respect to an unknown distribution $P_B(\{\mathbf{J}_i\})$, which we have to find. Whatever was codified into the prior distribution, it was for a reason. The new distribution will have to include the new information and in some sense, from all those that do, will have to "lie closer" to the prior. Closer means, effectively, that the fewest unwarranted new hypotheses must be introduced. The method to do this exists, and has its roots in Boltzmann, Gibbs, Shannon and Jaynes. See [15] for a modern exposition and justification of the Maximum Entropy method.

The resulting method consists of maximizing the cross entropy between the prior and the posterior distributions, subject to the constraints imposed by the new information and normalization. The constraints are included via the usual method of Lagrange multipliers:

$$S[P_B||P_0] = -\int \prod_i d\mu(\mathbf{J}_i) P_B \ln \frac{P_B}{P_0} + \alpha\left(E - \int \prod_i d\mu(\mathbf{J}_i)\mathcal{H}P_B\right)$$

$$+ \lambda\left(1 - \int \prod_i d\mu(\mathbf{J}_i) P_B\right), \tag{6}$$

with $d\mu(\mathbf{J}_i)$ being the uniform measure on the surface of a sphere in $M_D = 5$ dimensions.

It follows that the probability of configuration $\{\mathbf{J}_i\}$ is the Boltzmann distribution

$$\mathcal{P}(\{\mathbf{J}_i\}) \propto \exp[-\alpha\mathcal{H}(\{\mathbf{J}_i\})]. \tag{7}$$

The Lagrange multiplier α is still free and has to be chosen to impose that the average value of $\mathcal{H}$ is E. The informational content of E and α is, therefore, the same. Expected values of quantities of interest can be calculated for different values of α and of any parameters that enter in $\mathcal{H}$, such as δ. We name the new parameter α the *peer pressure*, since it sets the scale of the effect of social cost, and measures the inverse level of noise in the communication channel.

4. Data on Moral Foundations

Data consisted of five-dimensional score vectors with components in the interval $[0, 5]$ representing the relevance attributed to each moral foundation. Each vector was also labeled by the subject's self-declared political affiliation from p.a. = 1 (very liberal) to p.a. = 7 (very conservative) [31, 37].

Scores were extracted from Moral Foundations Questionnaires (MFQ30)[d] taken by $N = 14250$ US citizens. These questionnaires combine Studies 1 and 2 reported in [31] and are composed by two parts each with 15 sentences (3 for each foundation) plus one verification sentence.

In the first part subjects are asked the question: "*When you decide whether something is right or wrong, to what extent are the following considerations relevant to your thinking?*". Answers are given by scaling sentences of moral content from "*not at all relevant*" (score = 0) to "*extremely relevant*" (score = 5). In the second part, subjects scale sentences of a moral content from "*strongly disagree*" (score = 0) to "*strongly agree*" (score = 5). A moral vector component is then the average of 6 scores corresponding to a particular moral foundation. Moral vectors $\mathbf{J}_i$ are obtained by normalizing score vectors.

5. Results

We describe the aggregate behavior of the model and compare it to the aggregate behavior extracted from MFQ30 questionnaires. By introducing appropriate order parameters we can compare both systems, the set of subjects and the agents, in a semantic free manner. The correlation of political affiliation and different cognitive styles is established by first showing that different cognitive styles are associated to different distributions of moral values in the agent model and noticing that different sets of moral values are associated to different political affiliations in the MFQ30 data. A group of socially interacting members with a diversity of cognitive styles will therefore present a political spectrum. Groups of conservative agents show larger in-group coherence while groups of liberal agents adapt faster to changes in the issues under discussion.

We now discuss the statistical signatures that can be used to characterize the effective number of moral foundations of an agent. We compare them with equivalent signatures derived from the MFQ30 data.

Our main concern is the difference in the distribution of weights attributed to moral foundations by self-declared liberals and conservatives. Numerical simulation techniques, briefly discussed in the appendix, show that the model can have two qualitatively different regimes depending on the parameters δ and α (Fig. 2). For low δ and low α, the system is in a disordered state characterized by random correlations between the moral state vectors of the agents. Increasing either δ or α, a transition line can be crossed into a partially-ordered society. Now the agents are correlated to a symmetry breaking direction $\mathbf{Z}$, the *Zeitgeist vector*, which can be regarded as describing the cultural environment. The average agent is parallel to $\mathbf{Z}$. Now we reorient $\mathbf{Z}$, by rotating the frame of reference, so that its components are equal

[d] Available at moralfoundations.org (accessed on 7 February 2011).

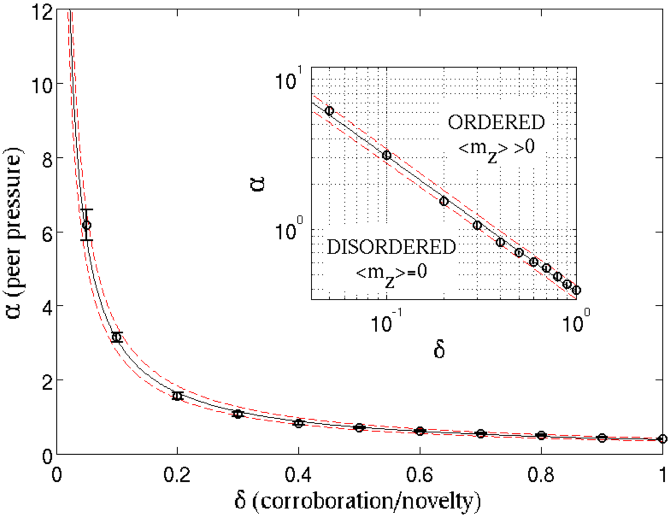

Fig. 2. The transition line separates phases with zero (below the line) and non-zero average overlap with the **Z** vector ($\langle m_{Z_i} \rangle > 0$). The phase transition is continuous. Symbols represent average and dispersion for 20 simulation runs of a $N = 400$ system with scale-free Barabási–Albert topologies. The full line represents a fit to the transition border line $\alpha = k/\delta$, with k constant. This can be seen more clearly in the inset. Dashed lines represent 95% confidence intervals for a regression.

(e.g. $1/\sqrt{5}$ each), explicitly assuming the equivalence of all moral dimensions. Note that opinions are rotation invariant and rotating makes no numerical difference. But it does foster interpretation, since a measure of the effective number of moral foundations of agent i can be defined as proportional to the sum over the moral dimensions a, of the agents moral weights:

$$m_{Z_i} = \sum_{a=1}^{5} J_{ia} Z_a, \tag{8}$$

the overlap between the moral vector and **Z**, ranging from -1 to 1. An agent with all moral dimensions equally important has $m_{Z_i} = 1$. Smaller values mean it relies on a reduced subset of moral dimensions. From the survey data, we extract, for each person a similar measure m_{Z_i} of their number of moral dimensions.

Our aim is to compare the statistics of m_{Z_i} from the data and from the model. Figure 3 compares histograms of m_{Z_i} as obtained from the data and as generated by the model, for $\alpha = 8$ on a scale-free social network. We have done several studies including different versions of the model. The conclusions we present, as far as the temptation of detailed quantitative confrontation with the data is tamed, are independent of the different variants of the model. We have only considered symmetric and homogeneous interactions which allows for the use of a single δ throughout the social network. In all simulations presented in this paper, the social neighborhood was represented by a scale-free random graph generated by a Barabási–Albert

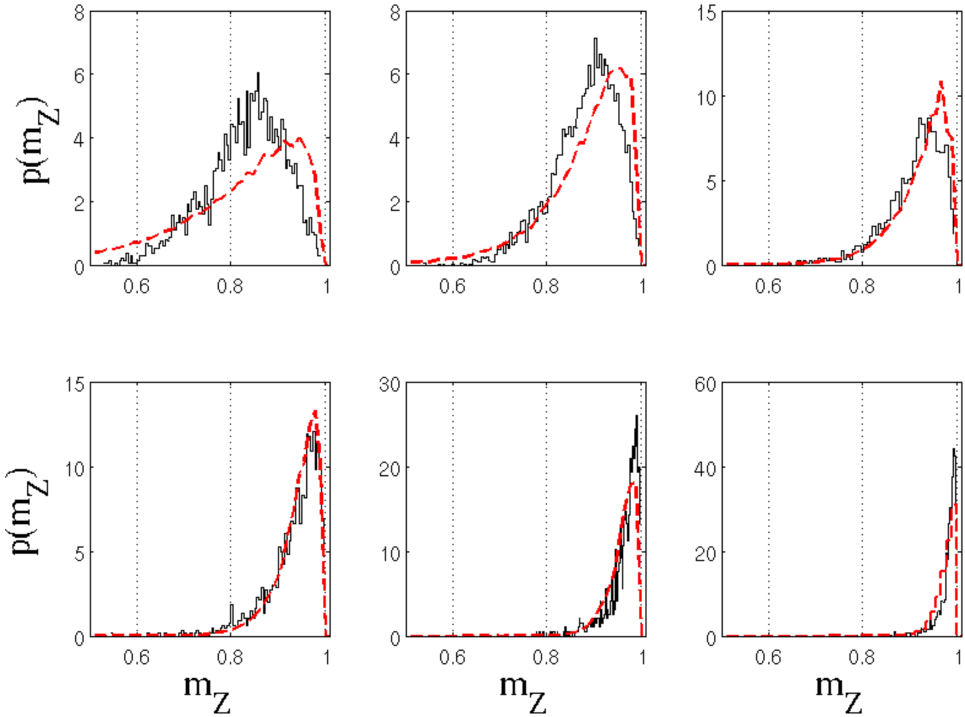

Fig. 3. Histograms for the effective dimension m_{Z_i} for p.a. = 1 to 7. Histograms for simulations at $\alpha = 8$ and different values of δ are depicted as dashed lines for comparison. Simulations have been performed with a scale-free social network of size $N = 400$. Results are qualitatively robust to changes in the lattice topology and system size.

model [46] with branching rate $m = 8$. We have also simulated a society subscribed to a two-dimensional square lattice with nearest-neighbor interactions. While this neighborhood seems too artificial, the results were qualitatively similar to those reported in this paper. Results within scale-free topologies, however, show the best agreement with data as far as overlap m_{Z_i} histograms depicted in Fig. 3 are concerned.

Agents have no political affiliation and persons do not declare their cognitive strategy δ. However, histograms permit identifying a political affiliation with a cognitive strategy. Figure 4 was prepared by calculating $\langle m_{Z_i} \rangle$ for each p.a. class of the data and then finding, for each given fixed α, the parameter δ that matches $\langle m_{Z_i} \rangle$ (20 Metropolis runs, $\alpha = 6$–12). Figure 4 also shows that the connection between the corroboration/novelty parameter δ and political affiliation is qualitatively robust for a reasonably wide range of α.

6. Discussion

The observations of the last section permit establishing the following link: political affiliations are partially derived from subsets of moral foundations, which arise

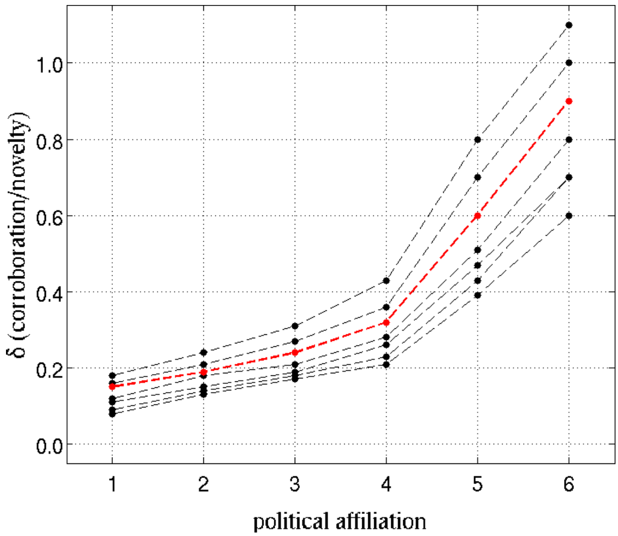

Fig. 4. By matching the mean average dimension, the relationship between the cognitive strategy parameter δ and the political affiliation is identified for $\alpha = 6$ to 12. The dashed line represents the case depicted in Fig. 3. Simulations were performed with a scale-free social network of size $N = 400$. Results are qualitatively robust to changes in the lattice topology and system size.

collectively from distinct cognitive strategies. We conclude that the link described in the literature connecting political affiliation to cognitive style [31, 36] arises as a consequence of social interactions.

As the order–disorder border line (Fig. 2) is approached from the ordered phase, the overlap with $\mathbf{Z}$ decreases, vanishing at the phase boundary. The best resemblance of the data and simulations occurs by identifying conservatives with agents far into the ordered phase and liberals with agents near the transition line but still in the ordered phase. Order and disorder refer to long range correlations and should not be attached to judgments of value.

We can go beyond the average number of moral foundations and make a prediction, based on the behavior of the agent model, about the width of the histograms. They decrease with increasing δ and the data shows that they decrease also with conservative tendency. The same identification: novelty seeking behavior to liberals, corroboration to conservatives, is again seen to arise as a consequence of collective behavior.

Order–disorder transitions can be driven by changing the peer pressure. Even without crossing the phase boundary, the model can be used to understand collective swings from left to right, as external conditions impose increased levels of peer pressure arising from the perception of threats. The reverse swing can also be understood when conditions demand higher adaptability to new challenges. We claim that with respect to moral issues, despite the differences in opinion derived from differential reliance on moral foundations, both conservatives and liberals are on the same side of the border. Other scenarios are discernible from the phase diagram. In an application outside the realm of morality, by looking at opinions on issues for which peer pressure might be lower, a group of large δ agents, relying on corroboration, could be found in a disordered phase and seem on this set of issues, to be liberal.

This theory is semantically neutral. Evolutionary considerations should be used to dress the theory with semantics and to understand why certain foundations of morality have emerged before others and why they are different, thus breaking the remaining symmetry between the five dimensions. Our model cannot claim to shed light on the different nature of the different moral foundations. It just states that based on differential treatment of novel and corroborating information, on conformity seeking behavior and on social interactions, populations will present collective statistically different moral valuations in a way that can be quantitatively described.

We believe that this work may create a number of opportunities for future research. Firstly, it is highly desirable to test the model against new data sets. We give three illustrative examples: (1) In [53], a connection is made between the structure of social networks during adolescence and political preferences in adulthood; (2) In [60], it is found that the opinion of a typical member of a virtual social network is influenced by about 20% of their neighbors; (3) [57] makes an empirical analysis of the voting patterns of US federal judge panels finding correlations between the political affiliation of the majority and the decisions reached. Our view is that the model we have proposed might have something to say about what sort of patterns are expected to be seen in each one of these examples.

We suggest that a social cost can be defined and that its mean value is directly associated to a parameter α we have identified as the "peer pressure". We regard the measurement of peer pressure as a relevant open problem suggested by our modeling effort.

We consider, however, as the most important contribution of this work to emphasize a particular methodological approach to the social sciences. From the description of how individuals react to incoming information obtained from social psychology empirical methods and neurocognitive data, we built an interacting model. Statistical mechanics leads to aggregated predictions which are tested against extensive data sets with partial information about populations. The exchange of information and the learning it elicits, induce collective emergent properties in the society not to be found in the individual. Presumably it may be useful to understand how cultural divides, such as those between conservatives and liberals, arise partly as consequences of diversity of neurocognitive mechanisms.

Acknowledgments

We thank Jonathan Haidt for kindly permitting us to have access to his data. This work received support from FAPESP and CNPq under grant 550981/07-1.

Appendix A. Methods Summary

A.1. *Metropolis sampling*

We assume that the model society is represented at micro-scales by continuous opinions h_j and that the system statistics can be described at intermediary time scales by a Boltzmann distribution:

$$P(\{\mathbf{J}_j\} \mid \mathbf{Z}) = \frac{1}{\mathcal{Z}(\alpha, \delta)} \exp\left[-\alpha \sum_{(i,j)} V_\delta(h_i, h_j)\right], \tag{A.1}$$

with (i, j) being edges of a social graph and $V_\delta(h_i, h_j)$ is the psychological cost given by (5). The statistics for the overlaps $m_{Zi} = \sum_{a=1}^{M_d} J_{ia} Z_a$, depicted in Fig. 3, can be obtained by sampling from (A.1). This has been done by employing a classical Metropolis sampling technique [47].

We choose a random *Zeitgeist* vector $\mathbf{Z}$. The distribution (A.1) is symmetric in relation to sign changes $Z_a \to -Z_a$ in the components of this vector. To deal with this degeneracy all simulations are started with moral vectors $\mathbf{J}_j$ aligned to the direction $\mathbf{Z}$. More realistic information exchange dynamics to be published elsewhere shows that the system auto-organizes into the same macrostates obtained by this simplified procedure.

A.2. *Wang–Landau algorithm*

While Metropolis-like algorithms sample from the distribution and collect data at a single point in the phase diagram, there is another class of algorithms which permit collecting information that will allow to obtain results for a set of parameter values. The Wang–Landau algorithm [63], belongs to this second class. The main theme is to collect information about the density of states, which in this case is peer pressure (α) independent and then to propagate to different values of α, by re-weighting via the Boltzmann factor. This is done for a particular value of the novelty/corroboration parameter δ. The density depends on δ and so this procedure has to be repeated for a set of δ values.

The transition line in Fig. 2 was obtained by Wang–Landau sampling of a system with Hamiltonian $\mathcal{H}$ at temperature $1/\alpha$ by finding numerically the maximum of the specific heat for fixed δ.

A.3. *Empirical histograms*

Data consisted of $N = 14250$ moral vectors with components related to five Moral Foundations in the interval $[0, 5]$ extracted from the MFQ30 questionnaire [31]. Each vector was labeled by the subject's self-declared political affiliation (from p.a. $= 1$ to 7). We first calculated normalized moral vectors $\mathbf{J}_i$ and, by defining the vector $\mathbf{Z}$ as the average vector within the conservative (p.a. $= 6$) and very conservative (p.a. $= 7$) classes, we have calculated histograms for the effective number of moral dimensions $m_{Zi} = \sum_{a=1}^{M_d} J_{ia} Z_a$ (depicted in Fig. 3).

Table 1. Moral foundations data: statistical summaries

p.a. *score*	n	$\langle m_z \rangle$	$\mu_{1/2}(m_z)$	σ_z	p.a. label
1	2919	0.825(5)	0.837(4)	0.084(2)	Very liberal
2	5604	0.877(2)	0.889(2)	0.069(2)	Liberal
3	2009	0.907(3)	0.920(3)	0.063(4)	Slightly liberal
4	1448	0.932(3)	0.947(3)	0.056(4)	Moderate
5	879	0.964(2)	0.975(2)	0.035(3)	Slightly conservative
6	1087	0.979(2)	0.986(1)	0.026(4)	Conservative
7	300	0.976(4)	0.987(2)	0.040(10)	Very conservative
$6+7$	1387	0.979(2)	0.987(1)	0.028(4)	Conservative

[a]Error bars represent 95% symmetrized bootstrap confidence intervals.
[b]$\mu_{1/2}(m_z)$ denotes the median of the overlaps m_z.
[c]We consider the classes "conservative" and "very conservative" together as their statistical moments, shown in the table above, are indistinguishable.

Data as well as an implementation in C of Metropolis and Wang–Landau sampling for the model we have proposed can be obtained in [16].

References

[1] Abrams, D., Wetherell, M., Cocharne, S., Hogg, M. A. and Turner, J. C., Knowing what to think by knowing who you are: Self-categorization and the nature of norm formation, conformity and group polarization, *Brit. J. Soc. Psychol.* **29** (1990) 97–119.

[2] Adorno, T., Frenkel-Brunswik, E., Levinson, D. J. and Sanford, R. N., *The Authoritarian Personality* (Norton & Company, 1969).

[3] Ahn, Y.-Y., Han, S., Kwak, H., Moon, S. and Jeong, H., Analysis of topological characteristics of huge online social networking services, in *Proceedings of the 16th International Conference on World Wide Web* (*WWW*) (2007).

[4] Aiello, W., Chung, F. and Lu, L., A random graph model for massive graphs, in *Proceedings of 32nd Annual ACM Symposium on Theory of Computing* (2000).

[5] Albert, R. and Barabasi, A.-L., Statistical mechanics of complex networks, *Rev. Mod. Phys.* **74** (2002) 47–97.

[6] Amodio, D. M., Jost, J. T., Master, S. L. and Yee, C. M., Neurocognitive correlates of liberalism and conservatism, *Nat. Neurosci.* **10** (2007) 1246–1247.

[7] Asch, S., *Social Psychology* (Prentice-Hall, 1952).

[8] Berns, G. S., Chappelow, J., Zink, C. F., Pagnoni, G., Martin-Skurski, M. E. and Richards, J., Neurobiological correlates of social conformity and independence during mental rotation, *Biol. Psychiat.* **58** (2005) 245–253.

[9] Borghesi, C. and Bouchaud, J.-P., Of songs and men: A model for multiple choice with herding, *Quanl. Quat.* **41** (2007) 557–568.

[10] Boyd, R. and Richerson, P. J., *The Origin and Evolution of Cultures* (Oxford University Press, 2005).

[11] Broad, C. D., *Five Types of Ethical Theory* (Routledge, New York, 2008).

[12] Campbell, L. and Garnett, W., *The Life of James Clerk Maxwell: With a Selection from His Correspondence and Occasional Writings and a Sketch of His Contributions to Science* (Macmillan, London, 1882).

[13] Casebeer, W. D., Moral cognition and its neural constituents, *Nat. Rev. Neurosci.* **4** (2003) 841–846.

[14] Castellano, C., Fortunato, S. and Loreto, V., Statistical physics of social dynamics, *Rev. Mod. Phys.* **81** (2009) 591–646.
[15] Caticha, A., Lectures on probability, entropy and statistical physics (2008), arXiv.org:0808.0012.
[16] Caticha, N. and Vicente, R. (2010), arXiv:1005.5718.
[17] Churchland, P., Toward a cognitive neurobiology of the moral virtues *Topoi* **17** (1998) 83–96.
[18] Dawes, C. T. and Fowler, J. H., Partisanship, voting, and the dopamine d2 receptor gene, *J. Polit.* **3** (2009) 1157–1171.
[19] de Waal, F., *Good Natured* (Harvard University Press, 1996).
[20] Durlauf, S., How can statistical mechanics contribute to social science? *Proc. Natl. Sci. USA* **96** (1999) 10582–10584.
[21] Eisenberger, N., Lieberman, M. and Williams, K., Does rejection hurt? An fMRI study of social exclusion, *Science* **302** (2003) 290–292.
[22] Engel, A. and Van den Broeck, C., *Statistical Mechanics of Learning* (Cambridge University Press, 2001).
[23] Forland, T. E., Mentality as a social emergent: Can the *Zeitgeist* have explanatory power? *Hist. Theory* **47** (2008) 44–56.
[24] Fowler, J. and Schreiber, D., Biology, politics, and the emerging science of human nature, *Science* **322** (2008) 912–914.
[25] Galam, S., Sociophysics: A mean behavior model, *J. Math. Sociol.* **9** (1982) 1–13.
[26] Galam, S., Majority rule, hierarchical structures and democratic totalitarism: A statistical approach, *J. Math. Psychol.* **30** (1986) 426–434.
[27] Galam, S., From 2000 BushGore to 2006 Italian elections: Voting at fifty-fifty and the contrarian effect, *Quanl. Quat.* **41** (2007) 579–589.
[28] Galam, S., Sociophysics: A review of Galam models, *Int. J. Mod. Phys. C* **9** (2008) 409–440.
[29] Galam, S. and Moscovici, S., Towards a theory of collective phenomena: Consensus and attitude changes in groups, *Eur. J. Soc. Psychol.* **21** (1991) 49–74.
[30] Gjoka, M., Kurant, M., Butts, C. T. and Markopoulou, A., Walking in facebook: Case study of unbiased sampling of OSNs, in *IEEE INFOCOM* (2010).
[31] Graham, J., Haidt, J. and Nosek, B., Liberals and conservatives rely on different sets of moral foundations, *J. Pers. Soc. Psychol.* **96** (2009) 1029–1046.
[32] Greene, J. D. and Haidt, J., How (and where) does moral judgment work? *Trends Cogn. Sci.* **8** (2002) 517–523.
[33] Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M. and Cohen, J. D., The neural bases of cognitive conflict and control in moral judgment, *Neuron* **44** (2004) 389–400.
[34] Greene, J. D., Sommerville, R. N., Nystrom, L. E., Darley, J. M. and Cohen, J. D., An fMRI investigation of emotional engagement in moral judgment, *Science* **293** (2001) 2105–2108.
[35] Haidt, J., The emotional dog and its rational tail: A social intuitionist approach to moral judgment, *Psychol. Rev.* **108** (2001) 814–834.
[36] Haidt, J., The new synthesis in moral psychology, *Science* **316** (2007) 998–1002.
[37] Haidt, J. and Graham, J., Planet of the durkheimians, where community, authority, and sacredness are foundations of morality, in *Social and Psychological Bases of Ideology and System Justification*, (eds.) Jost, J., Kay, A. and Thorisdottir, H. (Oxford University Press, 2009).
[38] Haidt, J. and Joseph, C., Intuitive ethics: How innately prepared intuitions generate culturally variable virtues, *Daedalus* (2004) 55–66.

[39] Holroyd, C. B. and Coles, M., The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity, *Psychol. Rev.* **109** (2002) 679–709.
[40] Katz, L. (ed.), *Evolutionary Origins of Morality* (Imprint Academic, 2002).
[41] Klucharev, V., Hytonen, K., Rijpkema, M., Smidts, A. and Fernandez, G., Reinforcement learning signal predicts social conformity, *Neuron* **61** (2009) 140–151.
[42] Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M. and Damasio, A., Damage to the prefrontal cortex increases utilitarian moral judgements, *Nature* **446** (2007) 908–911.
[43] Liljeros, F., Edling, C. R., Amaral, L. A. N., Stanley, H. E. and Aberg, Y., The web of human sexual contacts, *Nature* **411** (2001) 907–908.
[44] Marcus, G., *The Birth of the Mind* (Basic Books, 2004).
[45] Moll, J., de Oliveira-Souza, R., Eslinger, P., Bramati, I., Mourao-Miranda, J., Andreiolo, P. and Pessoa, L., The neural correlates of moral sensitivity: A fMRI investigation of basic and moral emotions, *J. Neurosci.* **22** (2002) 2730–2736.
[46] Newman, M., Barabasi, A.-L. and Watts, D., *The Structure and Dynamics of Networks* (Princeton University Press, 2006).
[47] Newman, M. and Barkema, G., *Monte Carlo Methods in Statistical Physics* (Oxford University Press, 1999).
[48] Pizarro, D. and Bloom, P., The intelligence of moral intuitions, *Psychol. Rev.* **110** (2003) 193–196.
[49] Quetelet, A., *A Treatise on Man and the Development of His Faculties* (W & R Chambers, Edinburgh, 1842).
[50] Roehner, B., *Driving Forces in Physical, Biological and Socio-Economic Phenomena* (Cambridge University Press, 2007).
[51] Schelling, T., *Micromotives and Macrobehavior* (WW Norton & Company, 1978).
[52] Schultz, W., Predictive reward signal of dopamine neurons, *J. Neurophysiol.* **80** (1998) 1–27.
[53] Settle, J. E., Dawes, C. T., Christakis, N. A. and Fowler, J. H., Friendships moderate an association between a dopamine gene variant and political ideology, *J. Polit.* **72** (2010) 1189–1198.
[54] Sherif, M., An experimental approach to the study of attitudes, *Sociometry* **1** (1937) 90–98.
[55] Shweder, R. A., Much, N., Park, L. and Mahapatra, M. M., The 'big three' of morality (autonomy, community, divinity) and the 'big three' explanations of suffering, in *Morality and Health*, (eds.) Brandt, A. and Rozin, P. (Routledge, New York, 1997).
[56] Somerville, L., Heatherton, T. and Kelley, W., Anterior cingulate cortex responds differentially to expectancy violation and social rejection, *Nat. Neurosci.* **9** (2006) 1007–1008.
[57] Sunstein, C., Schkade, D., Ellman, L. and Sawicki, A., *Are Judges Political?* (Brookings Institution Press, 2006).
[58] Sutton, R. and Barto, A., *Reinforcement Learning: An Introduction* (The MIT Press, 1998).
[59] Traud, A. L., Kelsic, E. D., Mucha, P. J. and Porter, M. A., Community structure in online collegiate social networks (2008), arXiv:0809.0960.
[60] Trusov, M., Bodapati, A. and Bucklin, R., Determining influential users in internet social networks, *J. Marketing Res.* **XLVII** (2010) 643–658.
[61] Vicente, R., Martins, A. and Caticha, N., Opinion dynamics of learning agents: Does seeking consensus lead to disagreement? *J. Stat. Mech-Theory* (2009) P03015.

[62] Vicente, R., Susemihl, A., de Andrade, J. P. J. and Caticha, N., A statistical mechanics approach for modelling social neuroscience data (2011).
[63] Wang, F. and Landau, D., Determining the density of states for classical statistical models: A random walk algorithm to produce a flat histogram, *Phys. Rev. E* **64** (2001) 056101.
[64] Weidlich, W., *Sociodynamics* (Dover, 2000).
[65] Woodward, J. and Allman, J., Moral intuition: Its neural substrates and normative significance, *J. Physiology — Paris* **101** (2007) 179–202.