




Article

Statistical Mechanics of Social Hierarchies: A Mathematical Model for the Evolution of Human Societal Structures

Nestor Caticha ¹, Rafael S. Calsaverini ² and Renato Vicente ^{2,*}¹ Instituto de Física, Universidade de São Paulo, São Paulo CEP 05508-090, SP, Brazil; ncaticha@usp.br² Department of Applied Mathematics, Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo CEP 05508-090, SP, Brazil; rafael.calsaverini@gmail.com

* Correspondence: rvicente@usp.br

Abstract: Social structure may have changed from hierarchical to egalitarian and back along the evolutionary line of humans. Within the tradition of sociophysics, we construct a mathematical model of a society of agents subject to competing cognitive and social navigation constraints and predict, using statistical mechanics methods, that its degree of hierarchy decreases with encephalization and increases with group size, hence suggesting human societies were driven from hierarchical to egalitarian structures by the encephalization during the last few million years and back to hierarchical due to fast demographic changes during the Neolithic. In addition, applied to a different problem, the theory leads to the following predictions for modern pre-literary humans: (i) an intermediate hierarchy degree in mild climates. In harsher climates, societies will be (ii) more egalitarian if organized in small groups (of less than 100 persons) but (iii) more hierarchical if in larger (of more than 1000 persons) groups. The predicted bifurcation, characteristic of a phase transition, is also seen in the empirical cross-cultural record (248 cultures in the Ethnographic Atlas).

Keywords: maximum entropy (MaxEnt); sociophysics; agent-based models; phase transition in societies



Citation: Caticha, N.; Calsaverini, R.S.; Vicente, R. Statistical Mechanics of Social Hierarchies: A Mathematical Model for the Evolution of Human Societal Structures. *Physics* **2024**, *6*, 629–644. <https://doi.org/10.3390/physics6020041>

Received: 31 October 2023

Revised: 11 January 2024

Accepted: 23 January 2024

Published: 19 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Behavioral phylogenetics makes it plausible that the common ancestor of *Homo* and *Pan* genera had a hierarchical social structure [1–5]. Paleolithic humans with a foraging lifestyle, however, had a largely egalitarian society, and yet hierarchical structures became again common in the Neolithic period. Today, nonliterate human societies fill the ethological spectrum [6] from egalitarian to authoritarian and despotic. This nonmonotonic journey U-shaped trajectory along the egalitarian–hierarchical spectrum during human evolution was stressed by Bruce M. Knauft [1]. Several anthropological explanations [1] have been constructed for Knauft’s U.

A number of recent papers [7–9] have focused on the evolutionary mechanisms that could resolve the apparent Darwinian paradox of the early-stage viability of hierarchical social structures. Since leadership positions provide preferential access to resources, it is not clear why followers would comply in the absence of coercive institutions. One potential answer is group-level selection [8] aimed at reducing scalar stress while enhancing in-group cooperation. This stream of study reveals some of the driving forces behind the connection between group size and the prevalence of hierarchy. Our work is intended to complement existing evolutionary arguments. In the context of physics, our main emphasis is on formulating an equation of state for the system, grounded in micromotives, as opposed to focusing on the processes that govern the temporal evolution of the system’s state.

Our approach to the study of social organization is based on an agent-based model and uses tools of information theory and statistical mechanics—a practice rooted in the tradition of sociophysics [10]. Sociophysics typically engages with mathematical tools to elucidate the interplay between variables in social systems. A path towards the mathematical analysis

of cross-cultural data is to proceed first by formulating a hypothesis, which might take the form where two given variables are related, either by virtue of their correlation or by some causal influence. Regression analysis from the data could then be used in the attempt to falsify the hypothesis. While an important tool, this method fails to suggest new relations that might appear from a more theory-based approach to model building. Instead, we model the society by considering agents that interact through the exchange of information. In particular, we are interested in the perception of each agent of the social network of its society, taking into account cognitive constraints and social navigation demands. Then, from general principles of information theory, we obtain the probability of the agent's perceived social network conditioned on the context. By calculating expected values of network node degrees, we obtain order parameters, which indicate whether the network is in an egalitarian-symmetric or hierarchical-broken symmetry state. The context depends on a combination of parameters, which grows with cognitive capacity of the agents and decreases with group size, modulated by environmental pressure. Since social perceptions mediate motivations and hence possible behaviors, we can make predictions about the effect of these variables in the probable forms of social organization and give a scenario for the Knauft's U-shape.

We built the model to address the U-shaped curve of egalitarianism. In itself, that is not surprising. Any experimental result can be explained by a model built explicitly for that purpose. But the same model permits making predictions about another empirical situation, dealing with the influence of ecology on the expected hierarchy of modern human groups. The theory suggests a form of looking at the available ethnographic data [11] and allows a new interpretation of observed patterns involving social structure, community size and environment in terms of a competition between cognitive constraints and social navigation demands, which are related to phase transitions. Furthermore, it allows for changes beyond the U-behavior to a zig-zag-like, such as that described in Ref. [12].

2. Model and Definitions

The central concept in this article is the cognitive representation of the network of social interactions by each agent, the inferred or perceived network and not the actual social network. Each one of the n political agents of a group has a perceived social network, represented by a graph of n vertices. For every agent i , the network is $S^i = \{s_{jk}^i\}$, where $s_{jk}^i = 1$, if agent i knows the social relation between agents j and k , independently of whether the relation is positive, neutral or negative. If agent i has no information about this relation, the bond is absent, $s_{jk}^i = 0$. This matrix is symmetric by its very design. For further details we suggest the following references [13–16] for models of the real social network with cognitive limitations and social interactions [17,18] for an agent-based model of how social stratification can emerge from efficiency demands of collective information processing. Reference [19] presents an evolutionary model for the emergence of inequity aversion as an adaptation selected from a pool of hierarchical groups. In addition, Refs. [7–9] treat other evolutionary approaches. The joint effect of demography and ecology on the emergence of hierarchy has been studied in Ref. [20]. Takao Terano and collaborators [15,16] used agent-based modeling and cognition-limited representations to analyze the emergence of money in terms of symmetry breaking in a doubly structural network representing a fixed social network and a dynamical network of exchangeability of commodities. We consider our paper as one that draws from and complements these earlier investigations.

Holding a particular belief about the network represents a cost to an agent. Our central hypothesis is that the cost can be represented by a function $C_0(S^i)$ arising from cognitive and social constraints. The cost is built by noting that:

- (1) The cognitive contribution to the cost increases with the number of memorized social relations N_{cog}^i . Then, the simplest form is

$$N_{\text{cog}}^i = \frac{1}{2} \sum_{j,k}^n s_{jk}^i. \tag{1}$$

- (2) It is preferable to have first-hand information about the nature of associations of third parties, compared to the cost of relying on indirect information derived from heuristics such as the ‘friend of a friend is a friend’ and its variations. Two agents j and k can be connected by a bond or a path of bonds, through intervening agents, so that their social relation can be estimated. Call l_{jk}^i the length of the shortest path of bonds joining j and k in the perceived network S^i . It is natural to assume that this lack of direct knowledge implies a social cost, which increases with the length of the shortest path between the agents.

Since S^i is the adjacency matrix, $M^i(\lambda)$ is defined by the integer powers λ of S^i (see, e.g., Ref. [21]):

$$M^i(\lambda) = S^{i\lambda}, \tag{2}$$

permit determining the shortest path

$$l_{jk}^i = \min \lambda, \quad \text{such that } M^i(\lambda)_{jk} > 0. \tag{3}$$

Define \bar{L}^i , the mean over all pairs (j, k) (counted only once)

$$\bar{L}^i = \frac{2}{N(N-1)} \sum_{j,k} l_{jk}^i. \tag{4}$$

Both N_{cog}^i and \bar{L}^i are functions of S^i , the variables of the problem.

We define the joint cognitive-social cost of the representation as the sum of the cognitive term and the social cost. The purpose of this exercise is not to uncover a ‘true’ form for the function that describes the cost to an agent. Our focus is on understanding the fundamental outcomes of the basic competition between costs rather than a precise identification of a functional form.

To be able to sum such different things, we introduce a parameter α :

$$C_0(S^i) = N_{\text{cog}}^i + \alpha \bar{L}^i. \tag{5}$$

The interpretation of the meaning of α comes from the following argument. For fixed values of N_{cog}^i and \bar{L}^i , as α increases, the relative contribution of the cognitive cost N_{cog}^i to the overall cost decreases. It is thus a proxy of the cognitive capacity of an agent. It can be assumed that α can be associated with the mass ratio between the neocortex and rest of the brain [22–24] and/or the medial frontal cortex including the anterior cingulate cortex [25–27]. The cost is infinite if there is no path between two agents in a perceived social network, so only connected graphs occur.

As this stands, an agent-perceived network is independent of the network of another agent. We now introduce interactions.

- (3) Many different forms of communication lead to social learning [28]. We follow ideas from opinion dynamics [29] and suppose that, through gossiping, agents share information or learn from social partners about coalitions of pairs of other agents, leading to correlated perceived networks. In the simulations, we introduce the possibility of an agent m copying from an interlocutor i its link about any other pair (j, k) .

Finally, we make a hypothesis that enters in the interpretation of the results but not in the construction of the model.

- (4) We suppose that perceptions about social coalitions engender motivations that direct behaviors. We call this the Perceptions Motivate Behaviors (PMB) hypothesis. Hence, there is a correlation between the perceived social network and the real coalitions emerging. We believe that this particular step of our argument should be amenable

to experimental verification. It receives support from the finding that in traditional societies, perceived positions in a social hierarchy are important to determine the commanded authority [30]. Within the PMB hypothesis, when the perceived social network is symmetric, social agents are not inclined to accept inequalities. Attempts to dominate a group, either economically, politically or sexually, would be met with resistance. However, if the symmetry is broken, social agents would not be so reluctant to accept a hierarchical political organization, at least under high rates of information exchange (or gossiping), when the central agent is with high probability the same for most of the agents.

3. The Probability of a Perceived Social Network

The standard methodology of information theory and statistical mechanics, which includes Bayesian and entropic methods, leads to the attribution of a probability $P(\{S^i\}|I)$ to a network S^i , conditioned on available information I .

Consider $E(C_0)$ the expected value of $C_0(S^i)$ under $P(S^i|I)$

$$E(C_0) = \sum_{S^i} C_0(S^i)P(S^i|I), \tag{6}$$

then the sum is over all 2^n possible configurations of S^i . Suppose that either the information available is that $E(C_0)$ has a specific value \mathcal{C} or that it has a significantly constant value during a relatively long time or alternatively, it is supposed that the value of $E(C_0)$ to be \mathcal{C} . In the last case, as \mathcal{C} is varied, different behaviors will be revealed.

The method of inference by maximum entropy introduces a parameter β . This parameter (or Lagrange multiplier) sets the range of fluctuations of the cost above its minimum value. Semantically, it is analogous to a pressure under which the band lives. Large values above the minimum cost are unlikely for large β but very likely if β is small. This pressure may arise from the ecology or it might model normative or informational peer pressure, from the perception of external threats, or as peer pressure arising from moral standards [31,32]. The name ‘pressure’ describes β in the generalized manner used in thermodynamics [33], in the sense that it is an entropic intensive parameter conjugated to the extensive cost in \mathcal{C} . In a simple fluid, the temperature is conjugated to the energy, but here, \mathcal{C} is not an energy nor β an inverse temperature.

The procedure calls for the maximization of the information entropy subject to the known constraints [34]. The entropy can be interpreted as a measure of the ignorance about S^i . Many probability distributions are compatible with the information constraint in Equation (6). Therefore, we take the one distribution that is most ignorant while still compatible with the information and the normalization requirement:

$$P(S^i|I) = \operatorname{argmax}_P \left\{ - \sum_{S^i} P(S^i) \log P(S^i) \right. \tag{7}$$

$$\left. - \beta_0 (\sum_{S^i} P(S^i) - 1) - \beta (\sum_{S^i} P(S^i) C_0(S^i) - \mathcal{C}) \right\}. \tag{8}$$

The result is the Boltzmann probability distribution

$$P(S^i|I) = \frac{1}{Z} e^{-\beta C_0(S^i)}, \tag{9}$$

where β , which controls the scale of fluctuations, is the intensive Lagrange multiplier conjugated to the extensive \mathcal{C} . The information content in the intensive β is equivalent to that in the extensive \mathcal{C} due to Equation (6) since they are conjugated variables. The partition function Z depends on the number of agents, α and β , $Z = Z(n, \alpha, \beta)$.

The macroscopic state of the system is characterized by the values of appropriate order parameters, expected values of functions of interest, which we now define. We probe

here whether represented agents are considered symmetrically or if distinctions are made. It is useful to consider the degree d_j^i of vertex j in a graph S^i , the number of agents k with $s_{jk}^i = 1$:

$$d_j^i = \sum_k s_{jk}^i. \tag{10}$$

In addition, we define the maximum degree and the average degree, respectively,

$$d_{\max}^i = \max_j d_j^i, \quad \text{and} \quad d_{\text{avg}}^i = \frac{1}{n} \sum_{j=1}^n d_j^i, \tag{11}$$

for agent i . Natural order parameters are the expectation values $E(d_{\max})$ and $E(d_{\text{avg}})$ with respect to $P(S^i|I)$.

We obtain estimates of the order parameters employing simple Metropolis Monte Carlo (MC) methods. We first considered isolated agents and did a MC simulation of the Boltzmann distribution (9). Calculations of the minimum distance were performed using Dijkstra’s algorithm. Qualitatively equivalent results can be obtained by defining a distance to be infinite if larger than a fixed finite length. Then, we simulate the exchange of information between the agents, which couples the different perceived networks.

At this point a methodological disclaimer is necessary. Our approach can be described as an application of maximum entropy inference, akin to the canonical statistical analysis of physical systems at equilibrium. The appearance of a Boltzmann distribution as a result of our method should not be interpreted as an excessive reliance on physical analogies. Rather, it demonstrates the extensive applicability of this inference technique.

4. Results

The expected values $E(d_{\max})$ of the maximum degree and $E(d_{\text{avg}})$ of the average degree, with respect to the probability distribution $P(S^i|I)$, are used to construct the order parameter H that signals hierarchical organization:

$$H = 2 - 2 \frac{E(d_{\text{avg}})}{E(d_{\max})}. \tag{12}$$

This is a helpful definition since for large $H \lesssim 2$, the symmetry of the perceived networks is broken and for small $H \gtrsim 0$, representations are symmetric. The factor 2 is included for convenience when comparing with the empirical record. A central result from the simulations is that, given β , to a good approximation, H does not depend on α and n independently but on the value of the combination $z = 2\alpha/(n(n - 1))$, the specific cognitive capacity per dyadic relation.

Our results can be described as follows, see Figure 1. For a given value of the pressure β , around 10, for sufficiently large specific cognitive capacity z (>1.2), H is close to zero. The typical perceived network of an agent is a fully connected graph (Figure 1, lower, on the right), meaning that cognitive costs are low enough and every possible pair of social interactions in the group can be stored. In the full graph, every agent is equal to any other. The representation of the network of social interactions is symmetric, and there is no agent that is seen as a central hub. The dynamics of social changes, a stochastic MC dynamics, is frozen.

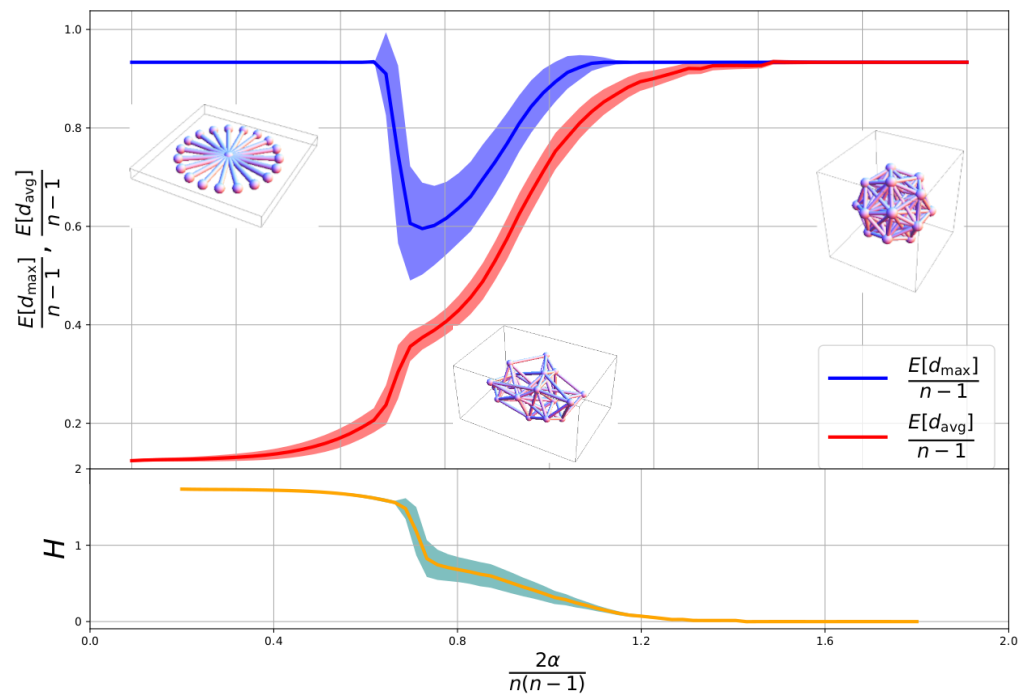


Figure 1. Upper: Monte Carlo estimates of the expected values of the maximum degree, averaged over the population, and the average degree of the social networks representation as a function of the specific cognitive capacity per dyadic relation, z , for $\beta = 10$. **Lower:** The order parameter H (12). Three regimes can be seen: for very large z , H goes to zero (symmetric phase) and almost all agents are equal. For very small z , H goes to 2, the broken symmetry phase. A particular agent occupies the central position of the network. An intermediate z transition region shows intermediate values of H . Error bars indicate the size of fluctuations obtained from the simulations. The insets are typical realizations of the inferred network of social interactions by an agent at that z position. Bonds are only drawn if the bond variable is present.

Either a smaller cognitive capacity or a larger size of the local social group, leading to intermediate z values, changes the above full graph picture (Figure 1, lower, in the center). Now, H is close to 1. Some bonds have to be absent in order to respect cognitive limits. The representations are still symmetric in a statistical sense, but fluctuations do occur, and a few agents will have a higher number of known social relations than others. The degrees of all vertices are not the same. Now, the dynamics of the perceived social network is not frozen, and different agents may temporarily occupy a more centralized position. To the perceiving agent, the members of its group are statistically equal, but its knowledge of their social connections is not uniform. Unequal roles start to appear in the inferred network, leading to the possibility of a social organization where there are ‘big men’ [35,36], and their distinctions might be described as those of first among equals, or rather among statistical equals in this theory.

For low z , resulting from an even larger number of agents or smaller cognitive capacity, a new regime for the organization of the inferred network emerges (Figure 1, bottom left), signaled by a larger value of $H \approx 2$. The symmetric role of the agents in the representation is broken. One of the represented agents has a special position, likely to be connected to most of the other agents, which in turn are connected almost uniquely to it. The dynamics again is frozen, and an agent crystallizes in a unique central position in the star topology of the inferred social network. Still, different agents have been sampled independently, and the central agents of the perceived networks are with high probability not the same.

Figure 2 shows the phase diagram, obtained by plotting the value of H in the (z, β) space. Figure 1 was obtained by a horizontal cut of Figure 2 at $\beta^{-1} = 0.1$.

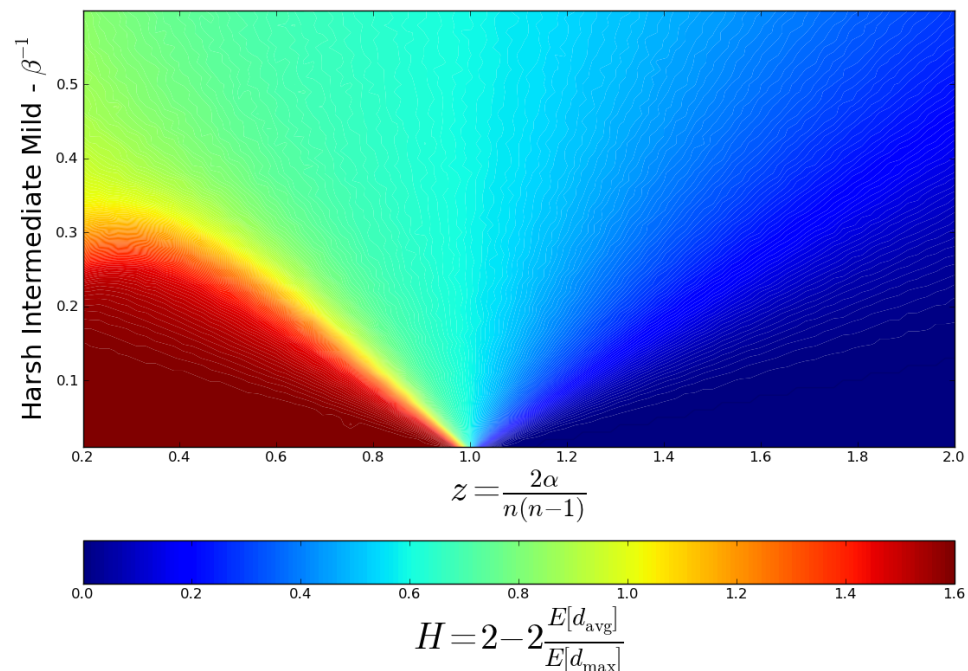


Figure 2. Hierarchical–egalitarian phase transition: phase diagram in the plane of specific cognitive capacity per dyadic relation, z , and the inverse ecological pressure, β^{-1} . The color code represents a measure of the social hierarchy measure or symmetry breaking parameter H . The red region is where the symmetry is broken (H near 2) and the maximum degree $E(d_{\text{max}})$ is much larger than the mean $E(d_{\text{avg}})$. The blue region is the unbroken symmetry phase, $H \approx 0$.

5. Correlation of Networks through Gossip

Gossip, though not a standard technical term in statistical mechanics, is employed here to denote the interaction between agents that arises through the exchange of information about interpersonal relationships. The impact of gossip is to create correlations among the perceived social networks of various agents. The three regimes—full graph, intermediate and star—remain unaffected.

We now run the simulation for the n agents together. A parameter g ($0 < g < 1$) measures the intensity of information exchange through gossip.

Choose an agent i and a pair (j, k) , independently of anything else, uniformly at random. A no gossip dynamic step occurs with probability $(1 - g)$: a MC Metropolis update is performed on the bond of the pair (j, k) . Let \bar{C}_0 be the joint cognitive-social cost with the bond s_{jk}^i replaced by \bar{s}_{jk}^i . With probability $\min(\exp(-\beta(\bar{C}_0 - C_0)), 1)$ let the change of s_{jk}^i by \bar{s}_{jk}^i be accepted. Otherwise, s_{jk}^i is kept fixed. With probability g , another agent l is chosen, also independently and uniformly at random and its corresponding edge s_{jk}^l is copied to \bar{s}_{jk}^i .

The step performed with probability $(1 - g)$ simulates the update of the social network representation by independent observations, learning new relations and forgetting about previously known relations. The gossip step, done with probability g , simulates the exchange of information where agent l tells and agent i learns or forgets something about the relation of agents j and k . Gossip can be introduced by more elaborate schemes, but this is sufficient for our modeling purposes.

After all i have been considered, a MC step has been completed. The α range $4.5 \leq \alpha \leq 90$ was divided uniformly into 100 intervals; the β range $0 < \beta \leq 20$ was divided into 200 intervals. The values of n varied from 7 to 15. We run a MC simulation for each fixed n and for each pair of α and β . 10^6 – 2×10^6 MC steps were run for thermalization, and then data about the order parameters were collected for around 4×10^6 MCS. It is usual in studying phase transitions to look at the large n limit of the population, since

singularities are only possible in the thermodynamic limit. However, since we are interested in the finite n case, we search for quantitative changes but not for critical exponents nor universality classes.

Figure 3 shows how frequent is the most frequent central agent as a function of the level of gossip g . Let $f_{ij} = P(c = j|i)$ be the probability that for the social network representation of agent i , the central element is j . The spreading of the probability distribution can be measured by the ratio of the average over the population of its entropy to the maximum possible value $\log n$:

$$\bar{s} = \frac{1}{\log n} \left(\frac{1}{n} \sum_i^n (- \sum_j f_{ij} \log f_{ij}) \right). \tag{13}$$

The results indicate that large correlation occurs when gossip dynamics dominates independent dynamics, starting around $g \approx 0.5$.

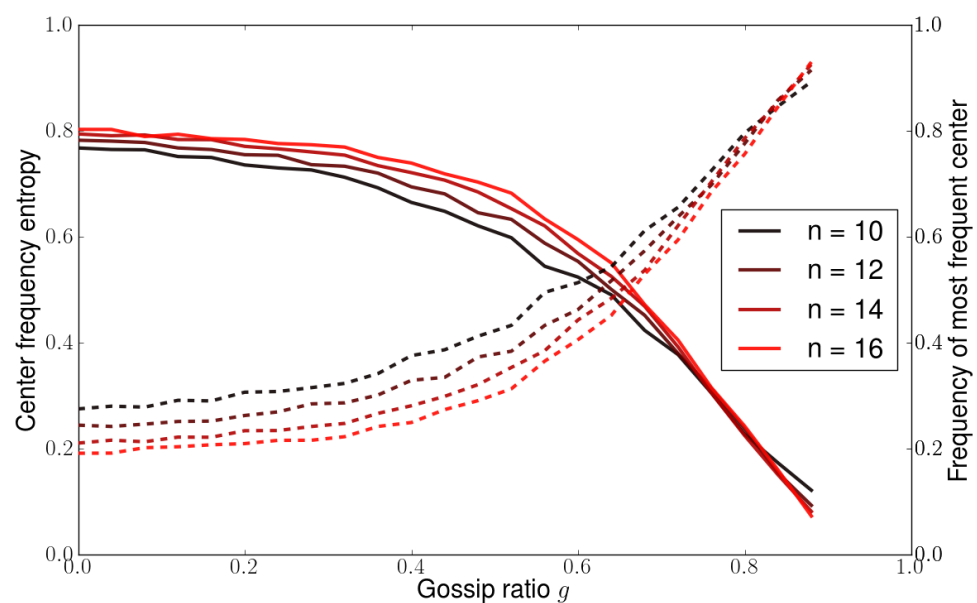


Figure 3. The effect of gossip inside the hierarchical phase. The entropy of the distribution of central agents (continuous curves) decreases with the increase of gossip, meaning that there is an onset of order where a particular agent preferentially occupies the centers of the stars. Every agent has a star network representation with a central agent. The agents share the same central agent with a probability that grows with g (dashed lines) showing another way of seeing the same ordering.

Figure 4 shows the acceptance rate for changes under the Metropolis dynamics in the network representations for three different β s. For high pressure, e.g., $\beta = 10$, and in the extreme cases of z , there is almost no tolerance to changes. In the large- z regime, upstarts that try to be different are not accepted. This is in accordance with the counter dominant behavior theory [2,5,37]. Perception of equality is hard to change, and this turns inequality unacceptable according to the PMB hypothesis.

In the low- z regime, the central agent is stable. Perceived network, highly correlated due to gossip, are very stable at both extremes of the scaling variable. The intermediate fluid phase decreases with increasing pressure β . Agents are only statistically equal and large dynamical fluctuations make the duration of the time intervals during which there is a change in the representational network much smaller than in the extreme egalitarian or hierarchical phases.

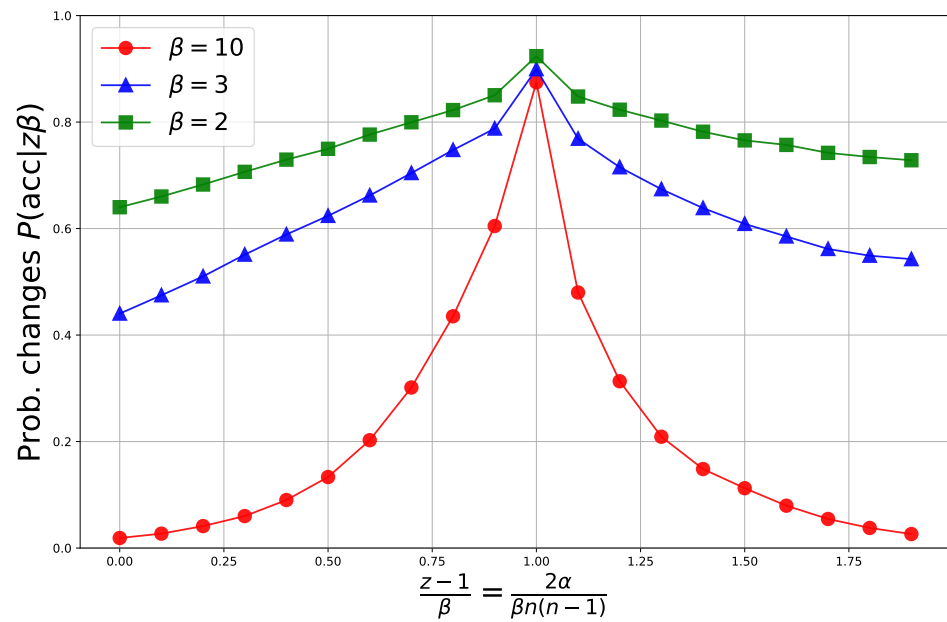


Figure 4. The probability of acceptance of changes $P(\text{acc}|z\beta)$ from the Monte Carlo simulations measures the tolerance to changes in the social network representations. As a function of z , for high pressure or harsher environments ($\beta = 10$), changes that would permit upstarts to be different from other agents are not tolerated for large z . This is analogous to counter dominance behavior theory [37]. Changes of the central agent of the star topology are unlikely to be accepted for small z . Milder pressures changes are accepted with high probability. Tolerance to changes is measured by the Monte Carlo acceptance probability.

6. Knauff’s U-Shape

The broken symmetry phase with the gossip-induced correlations, together with the PMB hypothesis, allows the possibility of appearance of hierarchical social structures.

Since the relevant scaling between α and n is through z , for either larger band size for fixed cognitive resources or small enough cognitive resources for a given band size, a fully symmetric representation of the network of social relations becomes impossible and an egalitarian organization of the group becomes unlikely.

A stylized scenario for the non-monotonic change along the hierarchy spectrum (Figure 5) follows from a historical reconstruction of the dynamics of conditions external to the model. A smaller cognitive capacity (or encephalization) in the evolutionary past of the agents, and hence a smaller z , is suggestive of hierarchical organization. Along the human lineage after it separates from that leading to chimpanzees and bonobos, there was an increase in the cerebral volume and in the neocortex. The reason for this is not under scrutiny. Whether social navigation, need for cooperative hunting or toolmaking were the sole or joint driving forces, the increase is factual. As cognitive capacity increases (larger z), the social representation can become a full graph and motivate an egalitarian organization.

On a faster time scale, a growth of the number of members of bands characterizes the Neolithic, bringing back the conditions (small z) for the symmetry breaking of the representation of the social network under fixed cognitive conditions. Again, the reasons for this growth are totally outside the scope of the present mathematical model, but we can only speculate that technological advances such as food production and storage might be involved [2,5,37].

The possibility of a further change of z on a still shorter time scale, restoring a symmetric representation of society, can occur if the means for processing more information are enhanced, even in the absence of cognitive changes through evolution, but rather through advances such as printing technology or even computer networking. Thus, the model not only suggests a mechanism for the U-shape trajectory but also to conceive further reorganizations of the social network representation and its consequences on social organi-

zation due to external conditions changing the balance between computational or cognitive resources and the size of the group.

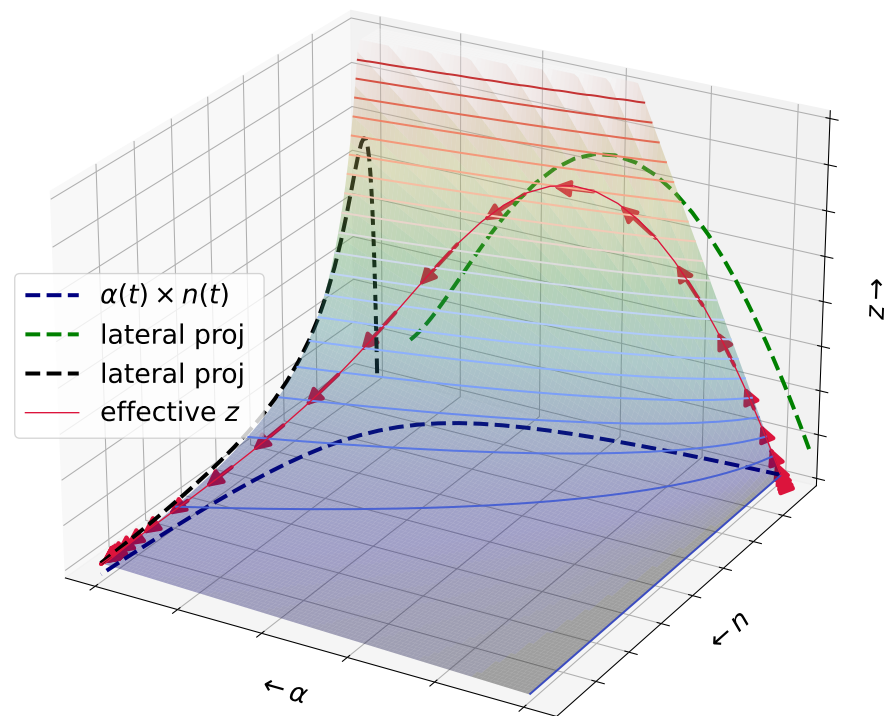


Figure 5. Schematic (inverted-) U-shape trajectory for the specific cognitive capacity, $z(\alpha, n) \propto 2\alpha / (n(n - 1))$, as a function of time (arrows). The higher the value of z , the more symmetrical or egalitarian the society will be. This is just a representation of externally caused changes in the cognitive capacity $\alpha(t)$ and the mean size, $n(t)$, of social groups as a function of time, showing a rapid increase of $\alpha(t)$ followed by an increase of $n(t)$. The dashed lines are the projections onto the respective planes. The contours are drawn for constant z values.

7. Social Organization: Theoretical Predictions and the Empirical Record

We could end our presentation with the suggestion that the reason for Kanuff’s U is the competition of navigational and cognitive costs. However, the model can be called to deal with a totally different problem. Can a signature of this competition between cognitive and social navigation constraints be seen today for modern humans? A clear theoretical prediction about the dependence of social stratification on ecological pressure β and group size ($z = z(n), \alpha$ fixed) can be compared to data from the ethnographic record. The prediction is divided into two parts. First, for very mild climates, that is, small ecological pressure β , intermediate social structures are expected. As climates of increasing harshness are considered, different social organizations will occur. Second, this difference depends on group size. Cultures organized in small groups will be more egalitarian, and those in larger groups more hierarchical.

7.1. Theory: Conditional Probabilities and Order Parameters

The phase diagram in the β - z plane can be thought of as a phase diagram in the β - n plane if the encephalization, or cognitive capacities, of the agents is fixed (see Figure 6). To be able to confront with the empirical record, we divided the ranges of β and z into three regions each: harsh, intermediate and mild climates and ‘small’, ‘medium’ and ‘large’ groups, respectively. The phase diagram is thus divided into nine regions. The regions are chosen essentially so that all points in the β - z space in the harsh-large region are of the same color (blue). The same is done for the region of harsh-small (all red) and for mild-small and mild-large. The qualitative results are robust to changes in the detailed choices.

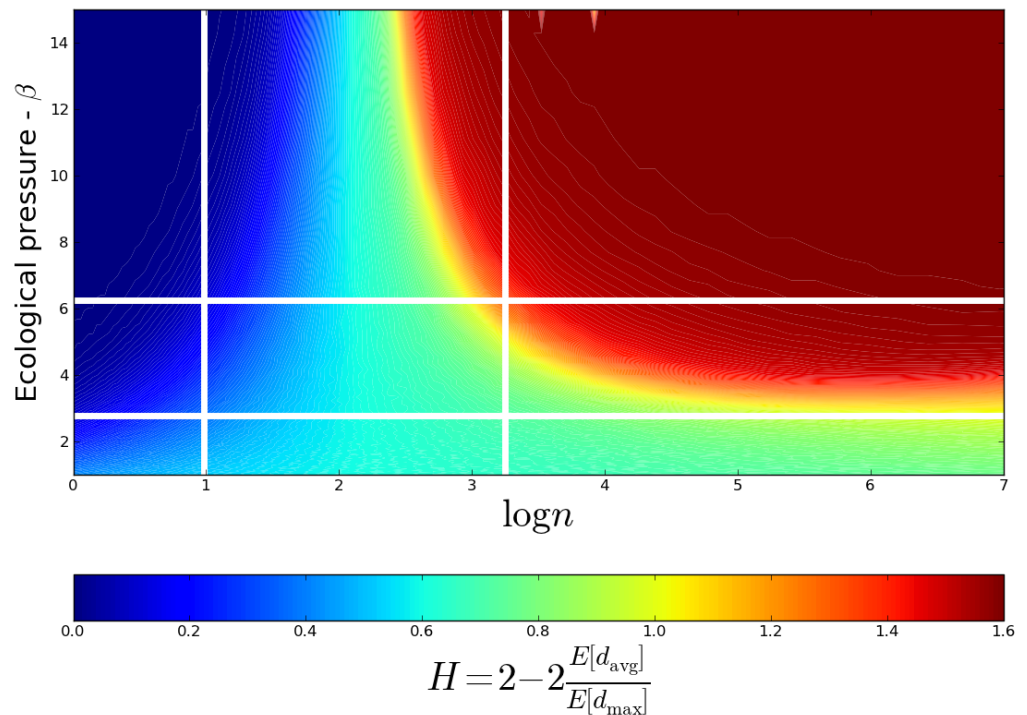


Figure 6. Phase diagram in the space of size of band versus ecological pressure, β . The color represents a measure of the social equality or symmetry H (12). The dark red region (large n) is where the symmetry is broken and the maximum degree is much larger than the mean. Dark blue (small n) is the symmetrical or egalitarian representation region. Separating those regions is a more fluid-like intermediate phase. The white lines show a choice of what is meant by large, intermediate and small both for β and n . The segmentation into nine regions is forced if the theory is to be compared to the empirical record.

A useful choice for the two values needed to separate the three regions are $\beta_H^{-1} = 0.15$, separating the harsh from the intermediate region and $\beta_M^{-1} = 0.35$ separating the intermediate from the mild region. For $\beta^{-1} > \beta_M^{-1}$, climate c is mild. For $\beta_H^{-1} < \beta^{-1} < \beta_M^{-1}$, climate c is intermediate, and for $\beta^{-1} < \beta_H^{-1}$, c is harsh.

For $z = \frac{2\alpha}{n(n-1)}$, the borders are set at $z_2 = 0.6$ and $z_1 = 1.4$ For $2 > z > z_1$, $s = 1$ small. For $z_2 > z > z_1$, $s = 2$ intermediate. For $0 < z < z_2$, $s = 3$ large. Then, we consider the mean order parameter $\bar{H}(s, c)$ for each region:

$$\bar{H}(s, c) = 2 - 2 \frac{\int_{\beta \in c} \int_{z \in s} D dz d\beta}{\int_{\beta \in c} \int_{z \in s} dz d\beta}, \tag{14}$$

where $D = \mathbf{E}(d_{\text{avg}}) / \mathbf{E}(d_{\text{max}})$.

7.2. Ethnographical Data

Data were obtained from the Ethnographic Atlas (EA) [11] using the Standard Cross Cultural code [38].

The relevant variables for our study are s , h and c , which stand for size category, hierarchy category and climate category. All variables can take integer values 1, 2 or 3. They are obtained by grouping the EA variables into three groups as shown in Table 1. There are 248 cultures in Ref. [11] with information on these three variables.

Table 1. Grouping of the Ethnographic Atlas (EA) variables [11,38] into categories. See text for details.

↓ Category/Value →	1	2	3
<i>s</i> : size (v31 ^a)	small	medium	large
<i>h</i> : social stratification (v66 ^a)	egalitarian	simple structure	complex
<i>c</i> : climate (v95 ^a)	harsh	intermediate	mild

^a The name of the variable in the EA.

7.3. Data: Conditional Probabilities and Order Parameters

The categorical values are obtained by grouping the relevant variables of the EA according to Tables given in the Supplementary Materials as shown in Table 1, into three categories.

We extract the numbers of cultures $N(s, h, c)$ with a given set of values (s, h, c) and the marginal numbers $N(s, c)$ of cultures with a given pair of values of (s, c) independently of h . These are related by $N(s, c) = \sum_{h=1,2,3} N(s, h, c)$. The conditional probabilities,

$$P(h|sc) = \frac{N(s, h, c)}{N(s, c)}, \tag{15}$$

of a culture having a given class stratification, given its climate and group size.

Then, we calculate the average hierarchy of the cultures with the same values of n and c , that is, that belong to the same size and climate categories. We calculate the empirical average hierarchies conditional on size and climate:

$$\bar{H} = \mathbb{E}(h - 1|sc) = \sum_{h=1,2,3} (h - 1)P(h|sc), \tag{16}$$

which satisfies $0 \leq \bar{H} \leq 2$.

Fluctuations around the average

$$\sigma_{EA}^2 = \mathbb{E}((h - \bar{h})^2|nc) = \sum_{h=1,2,3} (h - \bar{h})^2 P(h|nc), \tag{17}$$

can be calculated to define error bars (see Supplementary Materials).

7.4. Theory–Data Comparison

The empirical and theoretical predictions' order parameters are shown in Figure 7, left, and Figure 7, right, respectively. one can The typical bifurcation of a phase transition as different values of climate harshness are considered, occur in both the empirical data and the theoretical results. The qualitative agreement between theory and empirical record supports that our methodology is capable of suggesting new ways of looking at the available ethnographic records, which can now come under scrutiny by the community of quantitative ethnography dealing with cross-cultural studies.

That inequality rises with group size and that there are ecological factors involved has been previously considered [5,39–42], but not how the rise is modulated by ecological pressure nor the hypothesis that this is due to the competition of cognitive and social navigation needs and therefore the influence of climatic pressure on hierarchy can be reversed by demographics. It is not a coincidence that the transition between egalitarian and hierarchical regimes can be presented in the same language as phase transitions and critical phenomena in physical systems.

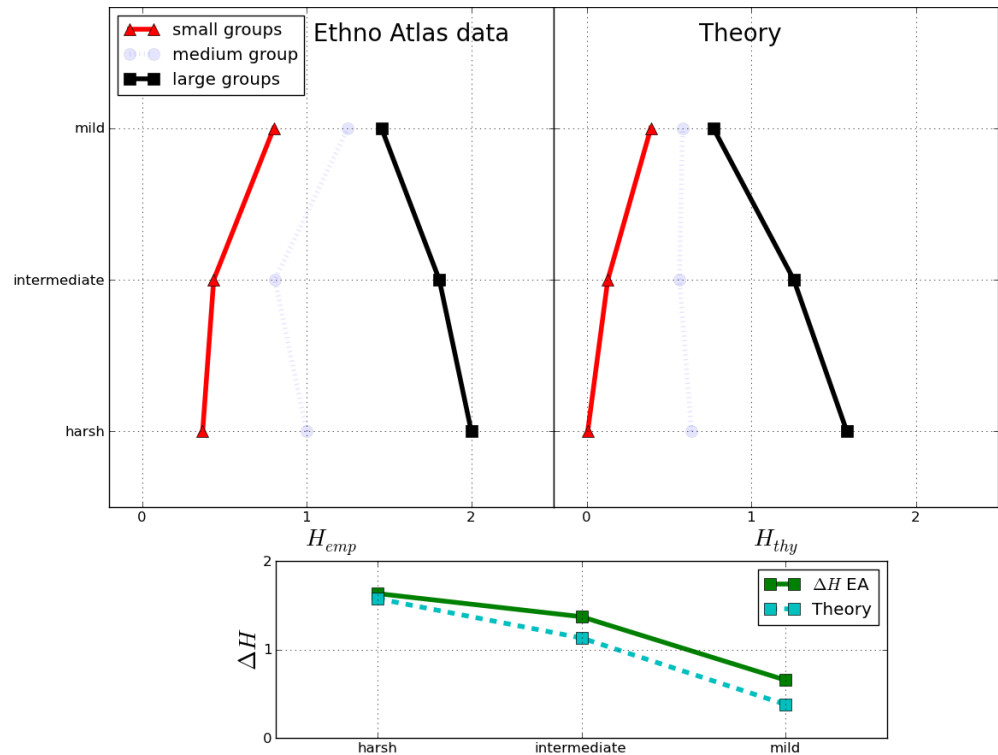


Figure 7. The bifurcation signature of the phase transition for Ethnographical Atlas (EA) data [11,38] (upper right) and theoretical prediction (14) (upper left). For mild climates, the expected hierarchies change with group size is quite small. For harsh climates, the expected hierarchy is larger for large groups and smaller for smaller groups. (Lower:) the difference in expected hierarchy ΔH between large and small groups decreases for milder climates. Harsh climates are tundra (northern areas), northern coniferous forest, high plateau steppe, desert (including arctic). Mild climates are temperate forest, temperate grasslands, Mediterranean, oases and certain restricted river valleys.

8. Discussion

Our approach, based on probability, entropic inference and information theory, is a methodological approach typical in the mathematical-physics modeling of systems that incorporates conditioning factors, however, they can be used when the conditioning comes from demographic, ecological, social and cognitive information.

Our main hypothesis is that a social-cognitive cost is relevant to characterize probabilistically the perceived social networks. The introduction of the conjugated parameter β , with the same informational content of the average cost, is an unavoidable theoretical consequence. It controls the size of fluctuations above the minimum possible value of the cost, prompting its interpretation as a generalized pressure. Gossip, a metaphor for information exchange, correlates the perceived networks. The cognitive capacity and the size of the group combine into a variable z , the specific cognitive capacity, and the perceived social state can be described in a space of just two dimensions (z, β) . External to the model, the dynamics of encephalization and band size determine the historical evolution of z , leading to a scenario for non-monotonic hierarchical change [1]. Further changes in z could occur, e.g., due to technological advances, which translate into more effective information processing and better social navigation. In addition, an effective reduction of ecological pressure, following enhanced productivity, can occur. Then, a more egalitarian perception of the social networks will follow. The PMB hypothesis predicts that motivations and behaviors will change, but the theory does not go into the area of predicting how behaviors change, nor what institutions will emerge in order to permit such behaviors, nor the time scales of these changes. Our approach to the transition from hierarchical to egalitarian and back dispenses the issue of whether the hierarchical type of behavior laid dormant (Lars Rodseth in Ref. [1]) and remained present throughout the Pleistocene or whether the

resurgence was due to convergent evolution (e.g., Ref. [43]). It can be turned on or off by the joint effects of cognitive resources, social demands, ecology and demography. These are similar to ordered–disordered transitions that result from pressure or temperature changes. The possibility of being solid ice is not dormant in water when it is heated up. At least that is not an appropriate metaphor, but it has been used in this context in the humanities.

One can speculate that the time spent in the large z egalitarian phase promoted conditions for the fixation of altruistic genes and the emergence of the ‘do unto others’ ideas since all are equal under the representation networks. One would not expect the fixation of altruistic behavior, which arises from punishment and collaboration [44–46] in other than the symmetric phase, but this should be amenable to model construction and analytic studies.

9. Conclusions

This simple model and the particular function we have used to represent the cognitive-social cost are far from complete. We do not claim specific numerical validation by confrontation with empirical data in any other way than just a qualitative one. More sophisticated forms of coalitions, other than dyadic pairing, should lead to increased richness of the phase diagram, without disrupting the rough overall picture. We have also avoided considering gender issues. Rampant sexual inequalities can exist in an egalitarian organization of males. Nevertheless, if competition between cognitive constraints and social navigation needs indeed occurs, then phase transitions from egalitarian to hierarchical perception follow from general arguments. It has been argued [47] that ‘in the history of the human species, there is no more significant transition than the emergence and institutionalization of inequality’. We expect that these methods, which unify the theoretical analysis of the empirical findings behind the scenario for the U-shape dynamics and the conditions that influence the transformation of perception of social organization, will stimulate the use of information theory methods in the analysis of empirical research in cross-cultural studies.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/physics6020041/s1>, Figure S1: Monte Carlo estimates of the maximum degree and the average degree of the social network representation as a function of $z = 2\alpha/n(n-1)$ for $\beta = 6$. This shows that the scaling of α with $n(n-1)/2$ is a good approximation that gets better for larger values of n ; Figure S2: Conditional Probability distributions $P(h|s, c)$ obtained from the Ethnographic Atlas data set; Figure S3: The standard deviation of the distribution of expected hierarchies, or range of values of hierarchy as a function of ecological pressure for small, medium and large groups. Left: From the Ethnographic Atlas data [38]. Right: Theoretical prediction; Table S1: The number of data points (ethnographic groups) in the EA relevant for this study; Table S2: The results for the empirical stratification \bar{H} ; Table S3: The results for the theoretical prediction H_T ; Table S4: V31. Mean Size of Local Communities. Table S5: V66. Class Stratification; Table S6: V95. Climate: Primary Environment (Coded by Frank Moore from Phillips’ Comparative Atlas); Table S7: EA Data: Cultures, Size, Class Stratification, Climate.

Author Contributions: Conceptualization and methodology, N.C. and R.V.; investigation, N.C. and R.S.C.; writing—original draft, N.C.; writing—review and editing, R.V.; visualization, R.S.C.; funding acquisition, N.C. and R.V. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially supported by Fapesp (2008/10830-2) and CNAIPS-USP, Brazil.

Data Availability Statement: Data is available from Supplementary Materials and https://growthecon.com/assets/Standard_Cross-Cultural_Sample-Codebook.pdf (accessed on 30 October 2023).

Acknowledgments: We thank Bruce Knauft for comments on the first stage of this project many years ago. While many of his comments led to changes in the presentation, many remain unaddressed and left for future work.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Knauff, B.M.; Alber, T.S.; Betzig, L.; Boehm, C.; Knox Dentan, R.; Kiefer, T.M.; Otterbein, K.F.; Apdock, J.; Rodseth, L. Violence and sociality in human evolution [and comments and replies]. *Curr. Anthropol.* **1991**, *32*, 391–428. Available online: <https://www.jstor.org/stable/2743815> (accessed on 20 January 2024). [CrossRef]
2. Boehm, C. *Hierarchy in the Forest: The Evolution of Egalitarian Behavior*; Harvard University Press: Cambridge, MA, USA, 1999. Available online: <https://www.jstor.org/stable/j.ctvjf9xr4> (accessed on 20 January 2024).
3. Boehm, C. Ancestral hierarchy and conflict. *Science* **2012**, *336*, 844–847. [CrossRef]
4. Dubreuil, B. Paleolithic public goods games: Why human culture and cooperation did not evolve in one step. *Biol. Philos.* **2010**, *25*, 53–73. [CrossRef]
5. Dubreuil, B. *Human Evolution and the Origins of Hierarchies. The State of Nature*; Cambridge University Press: New York, NY, USA, 2010. [CrossRef]
6. Vehrencamp, S.L. A model for the evolution of despotic versus egalitarian societies. *Anim. Behav.* **1983**, *31*, 667–682. [CrossRef]
7. Gavrillets, S.; Auerbach, J.; van Vugt, M. Convergence to consensus in heterogeneous groups and the emergence of informal leadership. *Sci. Rep.* **2016**, *6*, 29704. [CrossRef] [PubMed]
8. Perret, C.; Hart, E.; Powers, S.T. From disorganized equality to efficient hierarchy: How group size drives the evolution of hierarchy in human societies. *Proc. R. Soc. B Biol. Sci.* **2020**, *287*, 20200693. [CrossRef] [PubMed]
9. Lin, K.Y.; Schank, J.C. Small group size promotes more egalitarian societies as modeled by the hawk-dove game. *PLoS ONE* **2022**, *17*, e0279545. [CrossRef] [PubMed]
10. Galam, S. *Sociophysics: A Physicist's Modeling of Psycho-Political Phenomena*; Understanding Complex Systems; Springer Science+Business Media, LLC: New York, NY, USA, 2012. [CrossRef]
11. Murdock, G.P. *Ethnographic Atlas*; University of Pittsburgh Press: Pittsburgh, PA, USA, 1967. Available online: <https://digital.library.pitt.edu/islandora/object/pitt:31735057895306> (accessed on 20 January 2024).
12. Turchin, P. *Ultrasociety: How 10,000 Years of War Made Humans the Greatest Cooperators on Earth*; Beresta Books LLC: Chaplin, CT, USA, 2015. Available online: <https://www.pdfdrive.com/ultrasociety-how-10000-years-of-war-made-humans-the-greatest-cooperators-on-earth-e196524525.html> (accessed on 20 January 2024).
13. Killworth, P.D.; Bernard, H.R. A model of human group dynamics. *Soc. Sci. Res.* **1976**, *5*, 173–334. [CrossRef]
14. Bernard, H.R.; Killworth, P.D. Why are there no social physics? *J. Steward Anthropol. Soc.* **1979**, *11*, 33–38. Available online: <https://hrussellbernard.com/publications/> (accessed on 20 January 2024).
15. Yamadera, S. Examining the myth of money with agent-based modelling. In *Social Simulation: Technologies, Advances and New Discoveries*; Edmonds, B., Troitzsch, K.G., Hernández Iglesias, C., Eds.; Information Science Reference/IGI Global: Hershey, PA, USA, 2008; pp. 252–263. [CrossRef]
16. Kunigami, M.; Kobayashi, M.; Yamadera, S.; Terano, T. On emergence of money in self-organizing doubly structural network model. In *Agent-Based Approaches in Economic and Social Complex Systems V. Post-Proceedings of the AESCS International Workshop 2007*; Terano, T., Kita, H., Takahashi, S., Deguchi, H., Eds.; Springer: Tokyo, Japan, 2009; pp. 231–241. [CrossRef]
17. Dávid-Barrett, T.; Dunbar, R.I.M. Processing power limits social group size: Computational evidence for the cognitive costs of sociality. *Proc. R. Soc. B Biol. Sci.* **2013**, *280*, 20131151. [CrossRef]
18. Dávid-Barrett, T.; Dunbar, R.I.M. Social elites can emerge naturally when interaction in networks is restricted. *Behav. Ecol.* **2014**, *25*, 58–68. [CrossRef]
19. Gavrillets, S. On the evolutionary origins of the egalitarian syndrome. *Proc. Natl. Acad. Sci. USA* **2012**, *108*, 14069–14074. [CrossRef]
20. Powers, S.T.; Lehmann, L. An evolutionary model explaining the Neolithic transition from egalitarianism to leadership and despotism. *Proc. R. Soc. B Biol. Sci.* **2014**, *281*, 20141349. [CrossRef] [PubMed]
21. Barabási, A.-L.; Pósfai, M. *Network Science*; Cambridge University Press: Cambridge, UK, 2016. Available online: <http://networksciencebook.com> (accessed on 20 January 2024).
22. Sawaguchi, T.; Kudo, H. Neocortical development and social structure in primates. *Primates* **1990**, *31*, 283–289. [CrossRef]
23. Dunbar, R.I.M. Neocortex size as a constraint on group size in primates. *J. Hum. Evol.* **1992**, *22*, 469–493. [CrossRef]
24. Shultz, S.; Dunbar, R.I.M. Species differences in executive function correlate with hippocampus volume and neocortex ratio across nonhuman primates. *J. Compar. Psych.* **2010**, *124*, 252–260. [CrossRef] [PubMed]
25. Amodio, D.M.; Frith, C.D. Meeting of minds: The medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* **2006**, *7*, 268–277. [CrossRef] [PubMed]
26. Yoshida, K.; Saito, N.; Iriki, A.; Isoda, M. Social error monitoring in macaque frontal cortex. *Nat. Neurosci.* **2012**, *15*, 1307–1312. [CrossRef]
27. Allman, J.M.; Hakeem, A.; Erwin, J.M.; Nimchinsky, E.; Hof, P. Anterior cingulate cortex: The evolution of an interface between emotion and cognition. *Ann. N. Y. Acad. Sci.* **2001**, *935*, 107–117. [CrossRef]
28. Whiten, A. Primate culture and social learning. *Cogn. Sci.* **2000**, *24*, 477–508. [CrossRef]
29. Castellano, C.; Fortunato, S.; Loreto, V. Statistical physics of social dynamics. *Rev. Mod. Phys.* **2009**, *81*, 591–646. [CrossRef]
30. Earle, T. *How Chiefs Come to Power: The Political Economy in Prehistory*; Stanford University Press: Stanford, CA, USA, 1997. [CrossRef]

31. Caticha, N.; Vicente, R. Agent-based social psychology: From neurocognitive processes to social data. *Adv. Complex Syst.* **2011**, *14*, 711–731. [CrossRef]
32. Vicente, R.; Susemihl, A.; Jericó, J.P.; Caticha, N. Moral foundations in an interacting neural networks society: A statistical mechanics analysis. *Physica A Stat. Mech. Appl.* **2014**, *400*, 124–138. [CrossRef]
33. Callen, H.B. *Thermodynamics and an Introduction Thermostatistics*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1985. Available online: <https://archive.org/details/thermodynamicsan0000call/> (accessed on 20 January 2024).
34. Jaynes, E.T. *Probability Theory: The Logic of Science*; Cambridge University Press: New York, NY, USA, 2003. [CrossRef]
35. Sahlins, M.D. Poor man, rich man, big-man, chief: Political types in Melanesia and Polynesia. *Comp. Stud. Soc. Hist.* **1963**, *5*, 285–303. [CrossRef]
36. Lindstrom, L. ‘Big man’: A short terminological history. *Am. Anthropol.* **1981**, *83*, 900–905. [CrossRef]
37. Boehm, C.; Barclay, H.B.; Knox Dentan, R.; Dupre, M.-C.; Hill, J.D.; Kent, S.; Knauff, B.M.; Otterbein, K.F.; Rayner, S. Egalitarian society and reverse dominance hierarchy [and comments and reply]. *Curr. Anthropol.* **1993**, *34*, 227–254. Available online: <https://www.jstor.org/stable/2743665> (accessed on 20 January 2024). [CrossRef]
38. White, D.R.; Burton, M.; Divale, M.; Gray, P.; Kprotayev, A.; Khalturina, D. Standard Cross Cultural Sample: Codebook. 2011. Available online: https://growthecon.com/assets/Standard_Cross-Cultural_Sample-Codebook.pdf (accessed on 20 January 2024).
39. Carneiro, R.L. On the relationship between size of population and complexity of social organization. *Southwest. J. Anthropol.* **1967**, *23*, 234–243. [CrossRef]
40. Bingham, P.M. Human uniqueness: A general theory. *Quart. Rev. Biol.* **1999**, *74*, 133–169. [CrossRef]
41. Fletcher, R. *The Limits of Settlement Growth: A Theoretical Outline*; Cambridge University Press: Cambridge, UK, 2007.
42. Summers, K. The evolutionary ecology of despotism. *Evol. Hum. Behav.* **2005**, *26*, 106–135. [CrossRef]
43. Smail, D.L. *On Deep History and the Brain*; University of California Press: Oakland, CA, USA, 2008. Available online: <https://www.jstor.org/stable/10.1525/j.ctt1pp7mz> (accessed on 20 January 2024).
44. Boyd, R.; Richerson, P.J. The evolution of reciprocity in sizable groups. *J. Theor. Biol.* **1988**, *132*, 337–356. [CrossRef]
45. Boyd, R.; Gintis, H.; Bowles, S. Coordinated punishment of defectors sustains cooperation and can proliferate when rare. *Science* **2010**, *328*, 617–620. [CrossRef] [PubMed]
46. Schonmann, R.H.; Vicente, R.; Caticha, N. Altruism can proliferate through population viscosity despite high random gene flow. *PLoS ONE* **2013**, *8*, e72043. [CrossRef] [PubMed]
47. Feinman, G.M. The emergence of inequality. In *Foundations of Social Inequality*; Price, T.D., Feinman, G.M., Eds.; Springer Science+Business Media: New York, NY, USA, 1995; pp. 255–279. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.