Simulations show that shame drives social cohesion

Klaus Jaffe

Universidad Simón Bolivar
Apartado 89000, Caracas 1080, Venezuela
kjaffe@usb.ve

Abstract. Using the agent based simulation model Sociodynamica, I explore the effect of various intensities of feelings of shame in cementing social bonds that affect the aggregate health (average longevity of agents) and wealth (GDP) of two artificial societies: A primitive society with no division of labor, and a monetized society with division of labor for three different tasks. The results show that agents who feel shame are successful in colonizing virtual societies, but the effect of shame in achieving cooperation differs between both societies studied They results also show that small levels of shame are better in enhancing social cohesion than large costs imposed with altruistic punishment.

1. Introduction

Aristocles of Athens, better known as Plato (427-347 AC) and Protagoras of Abdera (c. 490 - c. 420 AC) recognized that feelings of shame are fundamental for the maintenance of social cohesion. We define shame as the instinct or innate drive that evokes a feeling of pain if one's behavior is disliked or rejected by others. Shame or equivalent social devices are detectable also among many animals, such as dogs, apes [36] and others. Shame is enhanced and fomented by modern and ancient cultures although exceptions to its usefulness for achieving social cohesion are known [25]. Its social power is used in the implementation of "bench marking" as practiced by business and management specialists, by using shame as a driving force for modifying behavior. As emotions can be simulated with artificial life [13], here we tried to make operational the concept 'shame' for simulation purpose. Operationalization of a emotion might be performed in many different ways [21], here we operationalize 'shame' by focusing on its social function, so that higher levels of shame increase the intensity of compliance with the social norm that elicited that feeling of shame. The aim of the exercise was to create a metaphor of a society that allows us to gain an increased understanding of the role of emotions in dynamic social interactions

One way to assess the effect of shame is in the context of reciprocity. Reciprocity is theoretically unlikely to evolve as a result of natural selection. Yet such behaviors evidently exist, and not only among humans [35]. No consensual explanation as to the forces allowing for its existence and maintenance exist [19, 31].

Thus human and animal cooperation and altruism continues to remain a puzzle. Among the most recent explanations proposed for solving this puzzle we might cite the following: Altruistic norms can 'hitchhike' on the general tendency of internal norms to be personally fitness-enhancing [32] and that a multi-level selection, gene-culture co-evolution argument then explains why individually fitness-reducing internal norms are likely to be pro-social as opposed to socially harmful [10].

Alternatively, neutral non-social players might stabilize the system enhancing the chances of altruistic behaviors to settle in social populations [12]. Another argument is based on the observation that reputation may foster social behavior among selfish agents, and is considerably more effective with punishment than with reward [31].

Previous studies have shown that behaviors that benefit others at one's own expense (altruism) are evolutionary stable (i.e. reach a Nash equilibrium) if many interactions are with genetically related [11] or with genetically similar individuals [14], so that kin selection or homophily can work. Also, repeated interactions allow for the recovery of the altruistic expense [1, 3, 23, for example], converting the act into social investment [15]. Simulations have been very important in this endeavor and interestingly, they suggest that societies allow for the occurrence of synergies that provide hidden or retarded benefits to the various actors, allowing natural selection to maintain behaviors that cement social bonding, such as altruism [14,15].

Evidence for long term hidden benefits of apparently altruistic behavior has been reported in the context of religions [2]. A more recent argument put forward to explain the maintenance of altruistic and of reciprocal behaviors is altruistic punishment. Support that the punishment of non-cooperators at a cost to the punisher and the punished may favor adaptation of altruistic behavior has been provided by field studies [9]. The agent based simulation model Sociodynamica confirmed that altruistic punishment is evolutionary stable, but due to its effect on the individual and society, should be rather viewed as a social investment [17]

Social simulations are one way to disentangle complex concepts. Such models have been used to study the effect of shame, although for different purposes, such as simulating "digital blush" [28]. The agent based simulation Sociodynamica [15,16] provides a framework that allows testing the robustness of social, psychological, biological and economic concepts. This model was applied to explore "altruistic punishment". Altruistic punishment means that individuals punish transgressors of fair play, although the punishment is costly for them and yields no direct material gain. Fehr & Gachter [9] proposed that altruistic punishment might explain the fact that people frequently cooperate with genetically unrelated strangers, often in large groups, with people they will never meet again, and when reputation gains are small or absent. These patterns of cooperation cannot be explained by the nepotistic motives associated with the evolutionary theory of kin selection and the selfish motives associated with signaling theory or the theory of reciprocal altruism. Fehr & Gachter [9] showed experimentally that the altruistic punishment of defectors is a key motive for the explanation of cooperation. Using human subjects, they showed that cooperation flourishes if altruistic punishment is possible, and breaks down if it is ruled out. They also showed evidence indicating that negative emotions towards defectors are the proximate mechanism behind altruistic punishment, confirming previous findings [29]. Yet, altruistic punishment or any other kind of social investment can not work if the individual punished does not respond to punishment. One such punishment is shaming, where short term economic cost may be involved but are not central to its implementation as other costs, such as future social opportunities, become more important. The feelings of the shamed agents may affect its future interactions with other agents.

In this paper I use the agent based simulation model Sociodynamica to explore the usefulness the concept "shame". I explore the effect of various intensities of feelings of shame on the aggregate health of a society with and without a rudimentary division of labor. The concept of shame, as used here, differs from that of image scoring [24] or shunning [26]. Here, the emotion is particular to the punished agent and determines its future behavior, not that of other agents. It is the reaction of the agents after receiving a punishment for not being sufficiently altruistic.

The results of this preliminary exploration of the effect of shame on virtual societies show that even if a reduced proportion of agents have even small levels of this feeling, the social regulation of the society is enhanced enormously.

2. Methods

The agent based computer simulation Sociodynamica was used to study the effect of altruistic punishment on aggregate health in artificial societies, measured as the average longevity of agents. A somewhat simpler version of Sociodynamica was described before (Jaffe 2002a, b). The model simulated a continuous two-dimensional toroidal world (500x400 pixels) through which different types of agents wander with Brownian motion, each at is proper speed. The speed of this motion (*m*) ranges from 0-30 pixels / time step. Agents did not learn. The simulations tested for the survival abilities of agents under variable circumstances. As dead agents were substituted by new ones, which had their parameters assigned at random, the simulations served as a way of weeding out those combinations of parameters that conferred low survival capabilities to agents, selecting those agents possessing parameters that conferred them larger survival possibilities. Agents did not inherit their parameters, as Sociodynamica is a metaphor for a society of agents living in a free competitive market with infinite replacement potential.

The virtual society was programmed to have different levels of interaction. A first level simulated interactions with the environment, where agents collected food and/or minerals. A second level simulated the interchange of goods between agents. A third level simulated agents that could show altruistic behavior by donating excess food to starving agents. A forth level simulated agents that punished stingy (non altruist) agents, with a cost to themselves and to the punished agent. The fifth level of simulated interactions included the feeling of shame in punished individuals, in which shameful agents would increase their altruistic donations after receiving punishment in accordance to their intrinsic level of shame.

Interactions with the environment

The toroidal world was supplied with patches of agricultural land (food resources: *Rf*) and mines (mineral resources: *Rm*). Each time an agent happened to land over one of these resources while walking randomly around, they acquired a single unit (*wo*) of the corresponding resource, accumulating wealth, either as food (*wf*) and/or as mineral wealth (*wm*). The probability of arriving and resting over a patch of food or minerals was dependent on the speed and type of movement of the agent. Similarly, its ability to serve as a trader also depended on these two parameters.

Agents spend some of their wealth in food in order to survive, consuming food at a basal constant rate (*b*), which was a fraction of the resource unit (*wo*). The wealth in food (*wf*) of each agent changed each time step:

$dwf = -b + wo$ where $wo = 0$ if no resources are encountered.

b determined the degree of external constraints or of competitiveness of the environment and was fixed at 0.1, indicating the speed of degradation of accumulated resources in *wo* / time-step. This value produced simulation outcomes that are closed to what we expect in real societies (Jaffe 2002a,b). Agents with no food resources left (*wf* = 0) perished and were substituted by a new agent with randomly assigned parameters. This substitution process allowed maintaining the total number of agents in the population constant.

Similarly, agents encountering minerals acquired a single unit of the resource (wo) each time they encountered it. Minerals never degraded (bm = 0). The wealth in minerals (wm) was inversely related to the probability of suffering a fatal accident for each agent. That is, mineral wealth improved the odds of surviving external "catastrophes" that killed agents at random, each time step, and large amounts of wm protected the agents against these catastrophes by reducing the probability of being affected by them. Agents with wm = 0 could survive, though, with a lower probability. The agents were struck by a fatal catastrophe if the following relation was true:

$wm < rnd(0\text{-}1) * D$

So that the greater the wealth of accumulated minerals of the agent, the lower their probability of being struck by a fatal accident or catastrophe, at any level of danger (D)

Both types of resources were replenished continuously. Each of them was concentrated in a different single patch and the total amount of resources was 200 *wo* for food and 100 *wo* for minerals. Each resource patch was distributed initially at random in the landscape but remained in the same place during the duration of each run.

Interchange of goods

Agents moved in random directions each time step. Each time an agent met another at a distance smaller than 20 pixels, an exchange of wealth could occur. These could be of two different types.

In some simulations, more "structured societies" were simulated by modeling labor specialization of the agents. In this case, agents were subdivided into three categories. Farmers which specialized in collecting only food; miners which collected only minerals; and traders. Traders specialized in trading minerals for food when encountering a farmer, and food for minerals when encountering a miner. These transactions were performed with money which allowed for the establishment of prices for commodities depending on their supply and demand. When not explicitly stated, artificial societies had no money and no division of labor: and all agents could collect food and/or minerals and could trade their goods.

Altruism

Donations of food occurred when the difference in food wealth (*wf1*-*wf2*) between the two agents was larger than 2. Then the richer agent transferred food to the less wealthy. The amount of food transferred depended on the generosity (*g*) of the donating agent, which varied initially among agents from 0 to 5 deciles of their wealth (*wf*), i.e. 0 to 50 % of their wealth.

As shown with Sociodynamica before, altruism can invade a population of agents if it provides social benefits, i.e. is rather more like a social investment [17]. Here we simulated that altruistic donation in food benefited society and the individual indirectly by increasing the *wm* of the donor in a percentage of the donation given by the variable *gain*, with a default value of 15%.

Shame

Agents were shamed by agents acting as altruistic punishers in the population. The proportion of shamers (altruistic punishers) was kept constant during each simulation run. The shamer imposed a fixed cost to the shamed agent and had to pay that same cost himself.

The intensity of the reaction of agents when shamed was coded as sensitivity to punishment or degree of the feeling of shame, which regulated the degree to which shamed agents increased their generosity after having been shamed. This increase could vary from 0 to 10 deciles of *wf*.

Simulation runs

The values of parameters not specifically analyzed or described as being fixed in a simulation, were allowed to vary randomly among agents in the ranges mentioned above. Unless stated otherwise, simulations were run 200 times with 1000 agents for 100 time steps. Although the configuration of populations never stabilized completely, after 60 time steps changes were very small. Thus, the populations of agents reached a stable state after time interval of 100 time steps.

3. Results and Discussion

The simulation results showed that the possession of high levels of shame was not a handicap for agents. Figure 1 shows the frequency distribution of agents possessing various levels of shame, after a simulation run of 200 time steps. Here simulations were run so that agents could adapt their levels for shame from 0 to 10 deciles, showed that they did not extinguish shame. That is, agents which felt no shame (shame = 0) did not displaced agents feeling shame. The figure shows the final distribution of agents (number of agents in the y axis) according to their levels of shame (in deciles of generosity increase after receiving a punishment given in the x axis) after a typical 200 time step run of a simulation of competing agents.

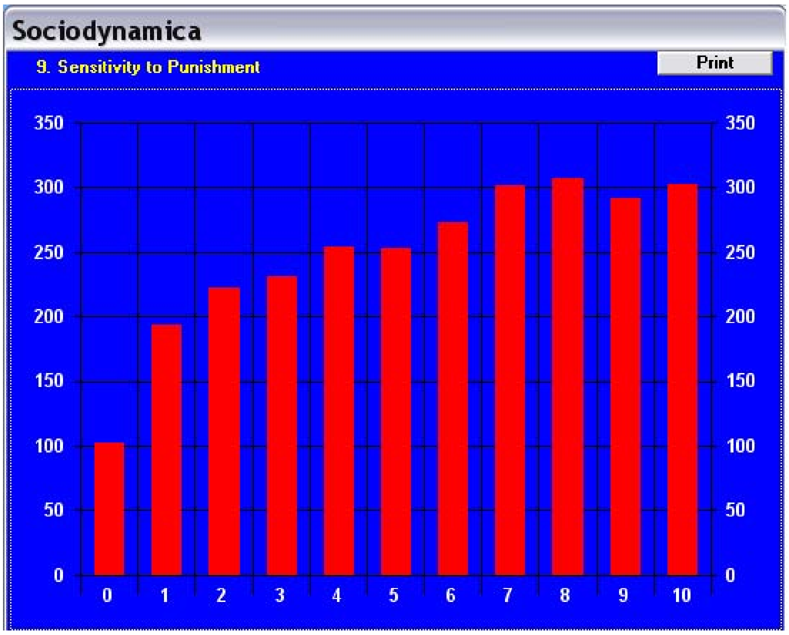


Figure 1: Distribution of levels of shame (sensitivity to punishment simulated as increased generosity shown after receiving a punishment) after 200 time steps in a virtual population of agents submitted to strong inter-agent competition.

Clearly, shame was adaptive in our model as increased generosity slightly increased the odds of surviving fatal accidents. For poor agents, their level of generosity had no effect on their fitness as their threshold for donating was never triggered.

As programmed in the model, levels of cooperation defined as amount of generosity, correlated with higher survival rates. This correlation, however, was non linear, as shown in the Figure 2. The results presented in Figure2, shows the average outcome of 200 simulations (each dot represents one simulation) after 100 time steps, regarding the levels of generosity (red squares, right axis) and longevity (blue circles, left axis) achieved by agents, with the programmed gain to society altruistic acts could achieve (gain through cooperation), as agents were living in worlds were cooperation could achieve different levels of social gains.

The figure shows a non linear relationship between the ethical variables (levels of generosity) and biological variables (longevity) of agents in relation to the amount of synergy (gain through cooperation) that the society could achieve through its altruistic behavior. The more society could gain through the altruistic actions of its agents, the more likely agents were altruistic, and the longer the average longevity of agents.

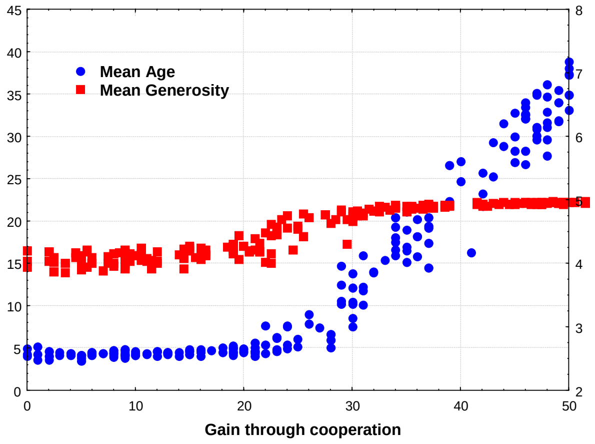
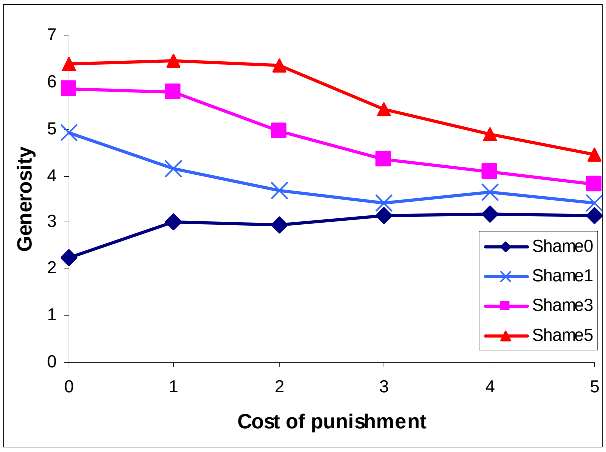


Figure 2. Result of 200 simulations (each dot represents one simulation) after 100 time steps, regarding the levels of generosity (red squares, right axis) and longevity (blue circles, left axis) achieved by agents.

In general, the interaction between the various parameters studied and its effect in increasing cooperation was complex and here we present only a selection of the relevant results.

In the next graphs (Figure 3 and 4) present the effect of the cost of altruistic punishment and that of the intensity of shame on the average age of the agents in the population. In the model, less cooperation increased the mortality of agents, thus, longer ages indicates that more successful cooperation was achieved. This effect was studied in primitive societies of barterers and collectors with all agents being generalists (no division of labor).

In these two graphs we see the effect of the cost of the altruistic punishment (paid by both, the punisher and the trader) on the average longevity and generosity of agents. The results show that the higher the cost, the lower the mean generosity of agents and the lower the life expectancy of agents, as lower levels of cooperation was achieved. We also see that a small cost reduces live expectancy significantly, but the effect in reducing the levels of generosity among the agents is much smaller. These effect were evident with all levels of shame explored.



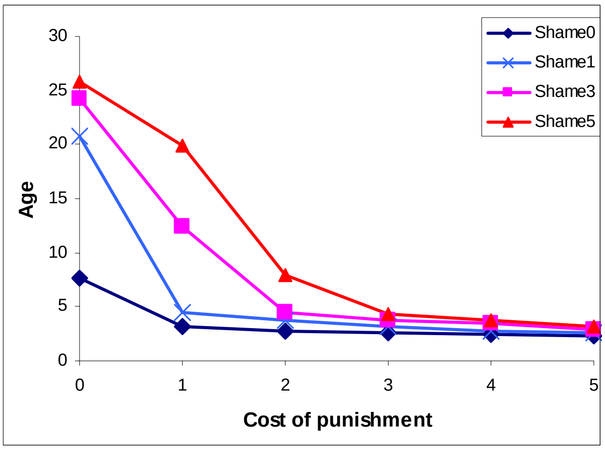


Figure 3 and 4: Average generosity and age of agents, after 200 time steps, for simulations with different costs of punishment and different fixed levels of shame.

The next two graphs (Figures 5 and 6) show the effect of increasing the degree of shame felt by agents that triggers increasing generosity, on the average longevity and generosity of agents. The results show an interesting non-linear effect. For large costs of altruistic punishment, no cooperation is established which reflects in low average longevity, whereas for smaller costs, small levels of shame allow the establishment of cooperation that achieves high longevity. With shame = 0 and costs = 0 we obtain cooperation that is not driven by punishment.

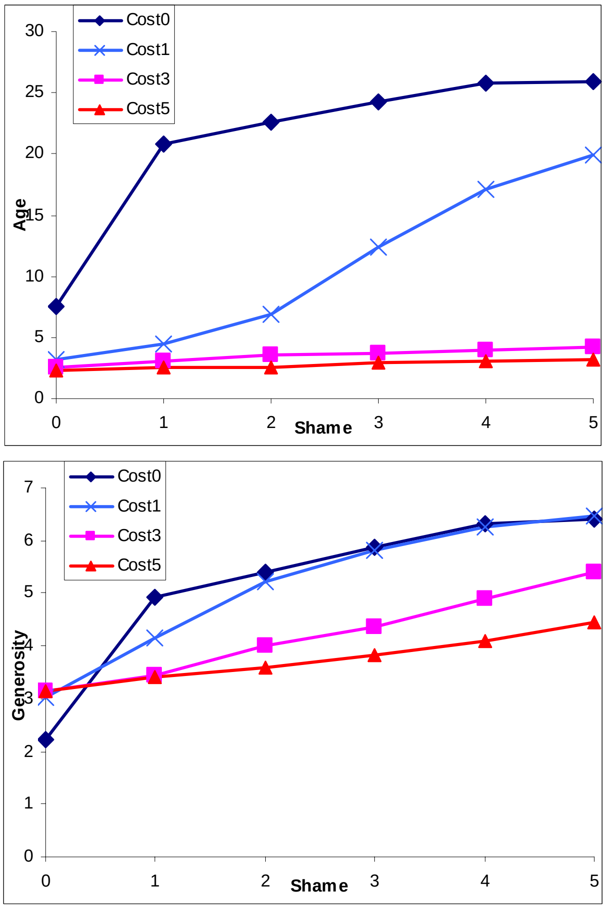


Figure 5 and 6: Average generosity and age of agents, after 200 time steps, for simulations with different costs of punishment and different fixed levels of shame.

All the simulations described until now were performed on artificial societies that had no money and no division of labor. But Sociodynamica also allows for the simulation of more complex societies. The next graph (Figure 6) shows the results of a simulation on a monetized society where agents were farmers, miners or traders (Division of labor).

Here interesting effects were evidenced. Although in general, the same trends as those described for simple societies were observed, these trends were less conspicuous, and a strange overlap of the curves was observed in the graph at low values of feelings of shame.

These results suggest that the effect of feelings of shame on some aggregate variable of a society will depends on the structure and relationship network of that society.

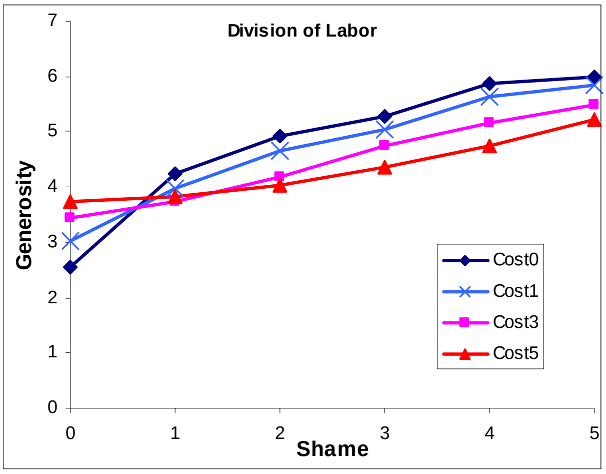


Figure 6: Relationship between generosity and shame for different costs of punishment in a more structured society.

4. Conclusions

The work is based on the assumption that cooperation is a social investment that eventually benefits the average agent, including many of the individuals doing the long term social investing or "altruistic" donation. In our simulations, the pro-social behavior simulated, benefited society as a whole, and being shamed had a cost that was reflected in the act that not being shamed increased the odds of survival in the future by small amounts. Under these conditions we saw an effect of both, the cost of altruistic punishment and the levels of shame of the agents, on the achievement of cooperation (level of generosity) in the virtual society. The interesting point is that both features, the cost of punishment and the reaction to punishment (shame), although part of the same behavior, differed in their effect on the aggregate properties of society. Large levels of shame favor social behavior whereas large costs for altruistic punishments hinder them.

The results clearly show that introducing shame in the virtual society is more effective for the achievement of pro-social behavior than increasing the punishment to non-complying agents.

In terms of evolutionary biology, shame might have been stabilized by evolution for any other purpose, as for example stabilizing family life and/or making education of offspring by parents more efficient. One the genetic and phenotypic tools for triggering shame have become well established, these could be used for other purposes, such as stabilizing societies. It is not the feeling of shame itself that is relevant for the adaptive behavior of a species but its role in triggering behaviors that stabilize or improve social relationships. Thus, shame can be viewed as a kind of long term indirect social investment, that helps the individual to regulates its behavior so as to improve the efficiency of social mechanisms that will eventually also benefit him.

This work shows the relevance of the context in which we analyze emotions. Looking at emotions in the wrong context will not unveil their adaptive function. Emotions that are exclusive to social species have to be analyzed in a social context in order to understand their working. Thus, much simpler models might not unveil features of emotions that are relevant to their social function. A more sophisticated analysis of social contexts might reveal novel functions for emotions such as shame.

Artificial societies and computer simulations, of course, are very distant from reality. The insight gained from this exercise, I suggest, should be that shame, and may be other feelings, do have important consequences on the aggregate properties of a society. The quantitative and qualitative distribution of feelings of shame in a real population should be explored in order to relate them to macro socio-economic variables of the society. This would allow correlating the prevalence and distribution of socially relevant emotions on the working of society. The results of the presented work would suggest that we might find some surprises when researching these questions in real societies.

References

1. Axelrod, R. and Hamilton, W. The evolution of cooperation. Science, 211: 1390–1396, 1981.
2. Berman, E. Hamas, Taliban and the Jewish Underground: An Economist´s View of Radical Religious Militias. NBR Working Paper 10004, 2003 http://www.nber.org/paoers/w10004
3. Boyd, R. and Lorberbaum, J.P.. No pure strategy is evolutionarily stable in the repeated Prisoner's Dilemma game. *Nature* 327, 58 – 59, 1987.
4. Boyd, R. and Richerson, P.J. Punishment allows the evolution of cooperation (or anything else) in sizable groups. Ethology and Sociobiology 13: 171-195, 1992.
5.Boyd, R., Gintis, H., Bowles, S., Richerson, P.J. The evolution of altruistic punishment. Proceedings of the National Academy of Science 100: 3531-3535, 2003.
6. Brosnan, S.F. and de Waal, F.B.M. Monkeys reject unequal pay. Nature 425: 297-299, 2003.
7.Castelfranchi, C. 2000. Engineering Social Order 1. ESA working paper, 2000. http://lia.deis.unibo.it/confs/ESAW00/pdf/ESAW04.pdf
8. Compte, A. 1830-1842. Cours de Philosophie Positive. Hermann Ed, Paris. 1998.
9. Fehr, E and Gächter, S. Altruistic punishment in humans, Nature 415: 137-140, 2002
10. Gintis, H. The hitchhiker’s guide to altruism: gene-culture coevolution, and the internalization of norms, Journal of Theoretical Biology 220: 407-418,2003.
11. Hamilton, W. D. The genetic evolution of social behavior. Journal of Theoretical Biology 7: 17-18, 1964.
12. Hauert, Ch., De Monte, S., Hofbauer J. and Sigmund, K. Volunteering as Red Queen Mechanism for Cooperation in Public Goods Game, Science 296, 1129-1132, 2002.
13. Hille, K. Synthesizing Emotional Behavior in a Simple Animated Character. Artificial Life 7: 303-313, 2001.
14. Jaffe, K. On the relative importance of Haplo-Diploidy, Assortative Mating and Social Synergy on the Evolutionary Emergence of Social Behavior. Acta Biotheoretica 49: 29-42, 2001.
15. Jaffe, K. An economic analysis of altruism: Who benefits from altruistic acts? Journal of Artificial Societies and Social Simulations 5: 3, 2002 http://jasss.soc.surrey.ac.uk/5/3/3.html
16. Jaffe, K. Monte Carlo exploration of mechanisms for the creation of aggregate wealth. Proceedings of IAREP/SABE Conference, Turku, Finland. 2002. http://atta.labb.usb.ve/Klaus/MonteCarlo%20Explo%20of%20Wealth.htm
17. Jaffe, K. Altruistic punishment or decentralized social investment? Acta Biotheoretica, 52: 155-172, 2004.
18. Jimenez-Alonso W. The role of kin selection theory on the explanation of biological altruism: a critical review. Journal of Comparative Biology 3: 1-14, 1998.
19. Johnson, D.D.P, Stopka, P. and Knights, S. The puzzle of human cooperation. Nature 421, 911–912, 2003.
20 Kaminka, G.A. and Tambe, M. Robust Agent Teams via Socially-Attentive Monitoring. Journal of Artificial Intelligence Rresearch 12: 105-147, 2000.
21. Krebs, D. and Janicki M. Biological foundations of moral norma. In: Psychological Foundations of Culture, M. Schaller and C. Crandal (Eds.), page 1-14.

22. Lincoln, R., Boxshall, G. and Clark, P. A Dictionary of Ecology, Evolution and Systematics. Second edition. Cambridge University Press. 361pp. 2001
23. Maynard Smith J. Group selection and kin selection. Nature 201: 1145-1147, 1964
24. Nowak, M.A. and Sigmund, K. Evolution of indirect reciprocity by image scoring. Nature 393: 573, 1998.
25. Orr S.W. The Economics of Shame in Work Groups: How Mutual Monitoring Can Decrease Cooperation in Teams. Kyklos, Volume 54, Number 1, 2001, pp. 49-66(18).
26. Panchanathan, K. and Boyd, R. Indirect reciprocity can stabilize cooperation without the second-order free rider problem Nature 432, 499-502, 2004
27. Pearce, D.W. (ed). The MIT Dictionary of Modern Economics. Forth edition. MIT Press, Cambridge Massachusetts. 474 pp. 1996.
28. Pitt, J. Digital blush: towards shame and embarrassment in multi-agent information trading applications. Cognition, Technology & Work. Springer London 6: 23 – 36, 2004.
29. Rawlings, E.I. Witnessing harm to other: a reassessment of the role of guilt in altruistic behavior. Journal of Personality and Social Psychology 10: 377-80, 1968.
30. Chimps Lack Charitable Nature. Nature 2005
31. Sigmund, K., Hauert, Ch. and Nowak, M. A. Reward and punishment, Proceedings of the National Academy of Science 98, 10757-10762, 2001.
32. Simon, H. A Mechanism for Social Selection and Successful Altruism, Science 250:1665–1668, 1990.
33. Smith, A. An Inquiry about the Nature and Causes of the Wealth of Nations. The Library of Economics and Liberty. 1776
34. Sober, E. and Wilson, D.S. Unto Others: The Evolution and Psychology of Unselfish Behavior. Harvard University Press, 700pp. 1999.
35. Wilson, E.O. Sociobiology: A New Synthesis, Harvard University Press. 1976.
36. Whiten, A. Horner, V. and de Waal, F.B.M. Conformity to cultural norms of tool use in chimpanzees. Nature 423: 737-740, 2005.