

Why Cooperate? The Place of Strong Reciprocity in the Evolution of Human Altruism

Zid Mancenido

Abstract

What makes humans cooperate is a core question of biology and anthropology. The question is no doubt complex but answers have been historically quite simple, revolving on singular categories of kinship or reciprocity. Recent findings in experimental economics have proposed a new answer: strong reciprocity, an empirically proven universal behavioural pattern that can account for the evolutionary development of human altruism. These proposals however have conceived it as quite a simple behavioural pattern rather than a complex suite of interrelated coevolved behaviours. This article critically examines the validity of strong reciprocity by considering the research methodology from which it developed and its component parts: altruistic punishment, conformist transition, third-party punishment, antisocial punishment, prosocial emotions, memory and trust, and life history and socialization. It argues that to answer the difficult question of cooperation, evolutionary models must be considered not just as coherent wholes but also disaggregated, with an attention to its coevolving component behaviours.

Introduction

It is generally accepted that the history of human evolution was complex and that human altruism has played a significant part in our ongoing survival as a species. There have been numerous models attempting to explain altruism (see Fehr & Fischbacher, 2003; Humphrey, 1997). This paper deals with a specific, relatively new evolutionary model of human cooperation termed *strong reciprocity*.

Four features define strong reciprocity. First, agents are predisposed towards cooperation. Second, agents will continue or increase their cooperation with those who cooperate with them. Third, agents will not cooperate with and will take extra costs to punish those who do not cooperate. And last, agents have a

tendency towards third-party sanctions against and punishment of those who do not cooperate. Simply put, it is a behaviour that rewards those who cooperate and abide by social norms and punishes those who do not. The costs associated with rewarding and punishing are borne by strong reciprocators even if this decreases their fitness relative to the group (Boyd, et al., 2003; Carpenter, et al., 2009; Bowles & Gintis, 2006; Fehr, Fischbacher, & Gächter, 2002; Fehr & Fischbacher, 2003, 2004; Gintis, 2000a; Gintis, et al., 2003, 2008).

The problem with current models of strong reciprocity is that they treat these four different features and their proximate determinants as a singular, relatively simple trait developed within a complex environment. This is done to differentiate it from and suggest its comparative advantage against other proposed evolutionary models of human cooperation such as kin altruism—where agents cooperate based on biological relations; reciprocal altruism—where agents cooperate because others cooperate with them; and by-product mutualism; where all agents benefit by contributing towards a public good. But strong reciprocity is neither simple nor singular: the behavioural ‘rules’ that undergird it aren’t just “ $rB - C > 0$ ” or “if you will help me, I will help you”. Rather, they are contingent on a complex of interrelated behaviours that are contingent on context-specific social norms and other-regarding preferences.

The following is an attempt to disaggregate the component parts of strong reciprocity in order to move forward in this debate over the evolutionary development of human cooperation. I begin with an assessment of the experimental economics games from which this evolutionary theory developed and the cultural variation amongst these results. I then turn to the component parts: looking at altruistic punishment, conformist transition, third-party punishment, antisocial punishment, prosocial emotions, memory and trust, and life history and socialization.

This paper is a response to current evolutionary models of strong reciprocity that speculate it could have evolved: from ‘a rather trivial modification of fitness-enhancing behaviours’ like kin altruism modified to include non-kin and/or reciprocal altruism modified to ignore future payoffs (Bowles and Gintis, 2004); or that it was only selected for through cultural group selection (Boyd, et al., 2003; Fehr & Henrich, 2004). By considering how the manifold cognitive and cultural elements that make up this uniquely human behaviour were developed and operate both individually and interactively, we can better understand that it was actually a complex of interrelated selection pressures that led to its evolutionary significance.

Experimental Economics Games

The biggest problem with models of strong reciprocity is that they are primarily derived from the games of experimental economics in which research participants are subjected to interactions and conditions that do not resemble hominin evolutionary history. Admittedly, these experimental economics games were not developed with evolutionary models in mind; they were produced to test specific behavioural patterns cross-culturally. Therefore, it is problematic to extrapolate evolutionary models from these findings without sufficient critical revision of research intent and design.

To explain, two commonly played economics games from which strong reciprocity has been derived are the Ultimatum Game and Public Goods Games. In the former, two players are allotted a sum of money. The first player offers a portion of the total sum to a second person. The second person can either accept or reject the first player's offer. If the second person accepts, she receives the amount offered and the proposer receives the remainder. If the responder rejects, then neither player receives anything. Public Goods Games are a little more complex in design but the premise is the same: interactions are based on resource allocation and cooperation between multiple people, each who have the opportunity to accept or reject cooperating in a group. Experiments are then modified slightly to test for particular variables and behavioural cues, including the addition of punishment, anonymity, repetition, and other stimulating factors. In brief, findings from experimental economics games such as these have shown that the picture of humans as rational maximizers is wrong; in no society has there been an Ultimatum Game with a mean offer near 0% (Bowles & Gintis, 2006; Fehr & Fischbacher, 2003; Gintis, 2000b; Henrich, et al., 2004, 2005).

I cannot dispute these findings but I have two serious methodological concerns. First, as Price (2008) argues, 'when subjects play economic games, their psychological adaptations for social behaviour are being deployed in environments that are radically different from those in which they evolved'. These highly structured experiments attempt to draw out universal behaviours by eliminating context. But evolutionary history was never modular; cooperative interactions were abundant and hominins were inundated with informational cues, such as reputation, relationships, and emotion (Sterelny, in press).

Frequent instances of cooperation and costly punishment of defectors in one-shot anonymous interactions have been used to support adaptive models that see cooperation as an end rather than a means (Fehr & Fischbacher, 2004; Gintis, et al., 2003; Henrich, et al., 2006). But this ignores the reality that historically most of our interactions were not anonymous, but were repeated, and our reputation was always at stake (Dreber, et al., 2008).

The response offered by Gintis (2011) and Fischbacher and Gächter (2010) is persuasive but does not deal directly with this problem. By considering first, how the rate of cooperation by players naturally “decays” as the games are repeated, and second, how the same players’ rate of cooperation is restored to its initial levels when they begin to play a second round of games (even after only a short break), Gintis argues that players completely understand the game situation. Cooperation decays within the game because players want to punish free-riders, not because they are learning how to maximize their payoffs in the context of highly structured economic experiments (as they would if the decay carried on to further sessions).

What this doesn’t deal with are the informational cues inherent to social environments. These findings show that, “The anonymity of the laboratory may indeed be sufficiently extraordinary that subjects simply play by the prudent and self-regarding rules of everyday life” (Gintis, 2010: 2), but that does not mean these rules are played within the *contexts* of everyday life. Anonymity and reputation aside, informational cues such as facial expressions, priming from previous experiences, the natural environment, or interested but uninvolved bystanders also affect the nature of human decision-making.

Second, experimental games only prove that people aren’t income-maximizing. This does not necessarily mean they are not fitness-maximizing. Price (2008) argues that strong reciprocity in its contemporary form is maladaptive; having evolved in conditions where strangers were rare and punishment hindered free riders’ development of fitness strategies. I don’t go so far, but I do question the link between income and fitness. It is true that fitness has always been measured by proxies; but some proxies, such as time, energy, and reproduction, are more cogent than others. By tagging cooperation to income, it includes quite contemporary attitudes about money; some of which are not fitness-maximizing. This is because “income” is a contemporary concept, born from the rise of capitalism and market structures that inhere values that did not exist during the Pleistocene Era. The influences of these attitudes on player behaviour need to be critically examined. Ultimately, evolutionary models can be inspired by contemporary behavioural experiments but these behaviours need to be given proper historical focus and mapped against more plausible evolutionary conditions.

Cultural Variation of Strong Reciprocity

If strong reciprocity is ‘a universal structure of human morality’ (Gintis, et al., 2008; Gintis, 2009) — and therefore governs altruistic behaviour — then we cannot ignore its phenotypic plasticity. We cannot ignore how it varies across

different human cultures. Early cross-cultural experiments suggested that variability wasn't an issue. Roth et al. (1991) found no significant difference in offers in Ultimatum Games played in Jerusalem, Ljubljana, Pittsburgh, and Tokyo. But more recent work attests otherwise; critiquing previous experiments for their engagement of too similar research subjects. Henrich (2000) presents Ultimatum Game results from the Machiguenga of the Peruvian Amazon, with average offers of 26 percent of the total: a far cry from the means of 40-50 percent observed until then in other experiments in Los Angeles and Yogyakarta (Indonesia) (Cameron, 1999; Roth et al., 1991). It is now generally accepted that culture greatly influences behaviour in economic games (Henrich, et al., 2004, 2005, 2006; Oosterbeek, 2004), especially in the extent to which strong reciprocators will go to punish those who defect (Gächter, Herrmann, & Thöni, 2010; Marlowe, et al., 2008; Wu, et al., 2009) and the occurrence of anti-social punishment (Gächter and Herrmann, 2008).

Cultural variability is important because strong reciprocity is contingent on social norms rather than on the more objective categories (such as biological relatedness or the concurrent or previous existence of cooperation) of other altruistic models. For strong reciprocity, cooperation and punishment are based on whether others adhere to or defect from expected culturally mandated behaviour. For example, Gurven et al., (2008) argue that one reason for the low average offers in economic games (32-37%) among Tsimane of the Bolivian Amazon is that they have 'no strong social norms governing a specific form or level of resource distribution'. The Tsimane still exhibit the behavioural pattern of strong reciprocity, just to a lesser, culturally determined degree.

This begs two questions about the development of strong reciprocity. First, if culture sets the bar for cooperative behaviour, what type of cultural norms 'set off' strong reciprocity? The term 'cooperation' is bandied about but there are many types, such as ecological, informational, reproductive. If the basic tenet of strong reciprocity is that violation of social norms encourages altruistic punishment, do these social norms include religion and ritual or alloparenting? Sterelny (in press) argues that cooperative behaviour coevolved between these complex social fields, as social interaction is not activity modular. Therefore, evolutionary models of strong reciprocity need to develop a more inclusive and robust framework to take these cultural and social structures into account.

Second, how closely did cultural norms and social institutions coevolve with strong reciprocity? Considering that the behaviour is contingent upon them, could it have stabilized with a small set of cooperative activities (such as hunting and reproduction) and then expanded to deal with other more complex ones characteristic of modern behaviours? Or was it the other way around? After all, it is not hard to imagine how the advent of social institutions in the Holocene would have selected for social conformity and regulatory behaviours akin to

the 'rule of law'. So far, studies on the variations in culture have been set aside in favour of the universalities of strong reciprocity, but consideration of the nature, limits, and guiding influences of this variability is necessary for the development of an evolutionary model.

Altruistic Punishment and Altruistic Cooperation

Strong reciprocators need to have the propensity, capability, and ability both to cooperate and to punish. Current models conceive these two as inextricable (cf. Gintis, 2000). This is true conceptually and phenotypically but we cannot assume that they mutually evolved from identical selection pressures and/or that their genetic biases are linked (West & Gardner, 2010). To illustrate the difficulties of this, Gardner & West (2010) model strong reciprocity as a greenbeard behaviour, where cooperation is a phenotypic marker indicating that the individual is also a punisher. Players are more likely to cooperate, considering once they see that someone is a cooperator, they understand this individual is also a punisher. Gardner & West argue that this sort of development is unlikely. This is because of the possibility of invading players who take advantage by mimicking the informational cues of greenbeards (to reap the benefits of others thinking they are greenbeards) but ultimately acting in a different way. In this case, Gardner & West argue that players who cooperate but do not punish or punish but do not cooperate are more likely to be dominant in the social environment.

Evolutionary models need to disaggregate these component parts of cooperation and punishment in order to gain a better understanding of their coevolved development. For example, structured teaching of cooperative behaviours cannot teach cooperation *and* punishment; but they can teach cooperation *or* punishment.¹ These behaviours are expressed differently and while they are contingent on the same informational processes to assess social norm compliance, the propensity, capacity, and ability to cooperate or punish utilize different cognitive and physiological processes. This is further complicated by the fact that cooperation and punishment are not zero-sum choices; one can choose to cooperate to a particular degree and choose to punish noncooperation to a particular degree. Experimental economics games express this by using scales in offers and punishments (Henrich et al., 2004).

So far it has only been established that insufficient cooperation can lead to punishment that is costly to the punisher. But there are other-regarding influences, such as the difficulty in punishing someone, the capacity for someone to punish, and the capacity for someone to defend themselves from punishment,

¹ However, Lehmann et al. (2006) argue that in small populations with little migration, behaviours of altruism and punishment could be identical in terms of resource availability and usage.

which affect the decision to cooperate or punish- some that are even mutually exclusive from one but influential on the other. For example, third party punishment occurs when a person punishes a defector in an interaction that they were not directly involved in. Third party punishment has been demonstrated in economic games but rarely to the same degree as second party punishment (Henrich, et al., 2006; Fehr & Fischbacher, 2004). What differentiates between these degrees? What other-regarding preferences and pressures determine the extent to which someone engages in costly punishment? These need to be taken into account because while some may be by-product adaptations, some may be key to the selection pressures that developed these behaviours.

Marlowe et al. (2008) find third party punishment much more prevalent in larger communities. In the evolutionary story then, is it possible that third-party punishment is a relatively recent addition to the complex of strong reciprocity, selected for in the Early Holocene as human groups became larger and there was a greater need for social control and stability? After all, costly punishment has been observed in other animal societies to coerce individuals and establish dominance (Clutton-Brock, 1995). Questions like these are brought to the fore when consideration is taken of the component parts of the whole behavioural suite. Because strong reciprocators don't just choose not to cooperate, but actively punish with costs to themselves, evolutionary models need to account for different coevolving developmental processes and selection pressures.

Costly Punishment and Conformist Transmission

It has long been assumed that costly punishment leads ultimately to conformism (Boyd & Richerson, 1985; Henrich & Boyd, 1998; Efferson et al., 2008) but recent experiments have shown that retaliation and non-conformism do occur (Denant-Boemont et al., 2007; Hopfensitz & Reuben, 2009; Nikiforakis, 2007; Wu et al., 2009). For example, and as will be explained in the next section, costly punishment can lead to feuds and other forms of anti-social punishment, such as those directed at players who cooperate *too much*. This disjuncture between theoretical models and empirical behaviour has centred on the questions: what gets people to conform, and what do they conform to? Theoretical models first suggested the fitness-enhancing benefits of conformism; not just at the individual level if the conformed-to behaviour was fitness-maximizing, but at a group-level if the behaviour reached fixation and for inter-group competition. But the propensity, capacity, and ability to conform is complex; Efferson et al. (2008) show that social learning is a behavioural tendency in and of itself, with some much better at identifying the fitness-maximizing behaviour and conforming more quickly as it becomes more common. While conformism can increase group homogeneity, accelerate fixation, and therefore be the

multiplier of strong reciprocity, greater consideration into the processes that people can 'acquire' the behaviour is necessary. Ultimately, the question must be, can coercion force strong reciprocity onto individuals weakened by lack of cooperation *and* punishment?

Current experiments signal no. For example, Nikiforakis (2007) presents evidence of retaliation following a quarter of costly punishments in Public Goods Games, attributing it to desire for reciprocal punishment and strategic motivations aimed at reducing future punishments. In evolutionary history, retaliation was most definitely a possibility; cooperation and defection were not managed by an economist; disagreements over cooperative norms such as food-sharing could most definitely have devolved into violence. Further, Hopfensitz & Reuben (2009) link retaliation to the social emotions of guilt and anger; the former triggers conformism and is activated when punishment (for one's lack of cooperation) is considered just; the latter triggers retaliation and is activated when punishment is considered unjust.

These suggest punishment and retaliation are expressions of personal opinions about social norm adherence and violation. Considerations of the evolutionary significance of costly punishment and conformist transmission have just begun. Future developments on evolutionary models of strong reciprocity need to take these into account. Conformism can be a fitness-enhancing behaviour, but the extent to which it was necessary for the spread of strong reciprocity needs to be critically revised. Conformism to the social norms and cultural practices of a dominant invading group may be easy for survival (cf. Boyd & Richerson, 2009), but acquiring (or even mimicking) the complex of strong reciprocity seems much harder.

Antisocial Punishment

Conformism is also contentious when costly punishment is not just limited to defectors. Herrmann, Thöni and Gächter (2008) report highly culturally variable instances of punishment directed at those who cooperate too much, even to the point that it inhibits the cooperation-enhancing effect of the punishment of defectors. Gächter & Herrmann (2008) extend these findings with one-shot interactions in four cities in Switzerland and Russia, isolating cultural differences in Russia as the root causes for more severe punishment of low contributors and antisocial punishment of high contributors. Their working hypothesis is that weak norms of civic cooperation and a weak rule of law are significant predictors of antisocial punishment.

Rand et al. (2010) model antisocial punishment as a mutated invading behaviour in a small structured environment with limited migration and show that it

can prevent costly punishment from enabling the evolution of cooperative behaviour. This is a clear problem for the efficacy of strong reciprocity. It is clear that antisocial punishment is culturally contingent; but it could be either evolutionarily tied to the bedrock psychology of strong reciprocity or a relatively recent adaptation of niche construction as a result of social institutions that rigidify fairness norms. Either way, it muddies the assessment criteria of cooperation, revealing complex relationships and contingencies in the evolution of strong reciprocity.

Prosocial Emotions

The Bowles & Gintis evolutionary model of human cooperation argues that prosocial emotions like shame evolved from repeated interactions of strong reciprocity as an informational shortcut to direct cooperation and punishment (2003 & 2006; cf. Gintis, 2004). It is true that emotions are proximate determinants of behaviour in economics games (and more generally, life). Harlé and Sanfey (2007) show that induced sadness results in higher offer expectations in the Ultimatum Game; and anger and guilt have been suggested as ultimate determinants of retaliation and punishment (Hopfensitz & Reuben, 2008). But the evolutionary story must be more complex; emotions direct much more than just cooperative behaviour; these other utilities were not by-product adaptations but coevolving design pressures. Further, the debate over prosocial emotions in non-human primates is unresolved; there is the possibility that shame is actually evolutionarily ancient. If this is true, it would validate Sterelny's argument for an inversion of the Bowles & Gintis model; specifically, that the bedrock psychologies of cooperation evolved first, and that strong reciprocity was dependent on cooperation, and therefore developed later (in press).

But emotions are not just psychological, they are physically displayed and used as information cues to influence social behaviour. The conditions under which human cooperation evolved must have been inundated with informational cues, and cooperation, defection, and punishment must have been physically effective. Considering emotions can be suppressed, manufactured, and faked, this suggests that these affective signals must have coevolved with either strong reciprocity or other informational processes — like a folk psychology — that could filter free-riding strategies like faking sadness or inducing guilt. While preliminary research has shown that rejection of unfair offers in Ultimatum Games is significantly lowered when responders can convey their feelings to the proposer concurrently with their decisions (Xiao & Houser, 2005), there is a way to go to test physically effective cues and their effect on cooperative behaviour.

Memory and Trust

Two underrated cognitive processes that are integral to strong reciprocity but underrepresented in the literature are memory and trust. In this paper it is proposed that they must have either coevolved, hitchhiked on other cognitive processes, or were hitchhiked upon, considering that strong reciprocity is so contingent on past interactions, assessments of social capability, and expectations about other people's social behaviour. For example, when partner choice in Ultimatum Games is allowed and previous results are public knowledge, more cooperative players are demanded (selected for) and reputational competition amongst players develops (Sylwester & Roberts, 2010). In this case, cooperation is a dynamic informational cue constructed through remembered past interactions that signal trustworthiness and therefore enable trust. Whatever the evolutionary dynamic, models of strong reciprocity need to take into account the extent to which the behaviour is reliant on memory and trust.

Reputation is the combined expression of these two processes of memory and trust and has been shown to significantly improve cooperation in experimental economics games (Bolton, et al., 2005; Brandt, et al., 2003; King-Casas, et al., 2005; Kreps & Wilson, 1982; Milinski, et al., 2002). In the conditions under which cooperation must have evolved, recognizing and internalizing informational cues about reputation must have been fitness-maximizing; reputation assessment would mean avoiding defection and maximizing opportunities for cooperation. These processes of acquiring, storing, and utilizing information must be considered as separate to (but of course, mutually evolving) with altruistic cooperation and punishment (Sterelny, in press). There is much written on the evolutionary development of memory; it has been integral to a whole suite of human faculties, including language and sensory functions. This literature needs to be accounted for within models of strong reciprocity. After all, just because these behaviours and processes are closely related, it does not mean they evolved simultaneously.

Life History and Socialization

So far, we have assumed that strong reciprocity is a universal behaviour based on studies of fully socialized persons. But developmental evidence offers a different picture, suggesting differences in cooperative behaviour based on life history stages (Harbaugh, et al., 2002; Sutter & Kocher, 2007). Nearly half of Grade Two students examined playing Dictator Games — similar to the Ultimatum Game, but without the ability for the responder to reject offers — gave nothing to their second players. This is a rationally maximizing behaviour, not replicated

in any other social group. Findings have also shown that the rate of cooperation in repeated interactions increases linearly from early childhood until the mid-twenties. These results support the intuitive picture of the phenotypic flexibility of strong reciprocity. After all, the behaviour of strong reciprocity is dependent on individuals internalizing particular social norms to learn when to cooperate or punish.

Developmental psychology has already provided evidence supporting the theory that children have a predisposition towards acquiring the social norms that make up strong reciprocity. Children understand what norms are and that they should be learned and followed: for example, those as young as two assume the world is full of social norms (Rakoczy, et al., 2008) and five-year olds can distinguish conventional rules that can be broken from natural laws that cannot (Kalish, 1998).

These studies bring to the fore two important questions in the development of strong reciprocity. First, *how do children internalize social norms?* As noted above, conformism is a possible but not a universal mode of transmission of the social norms that make up strong reciprocity. Sterelny (in press) offers an evolutionary story of niche constructed socialization beginning from tolerated watching to intentional teaching. If this true, the social norms upon which strong reciprocity is contingent must somehow have developed piecemeal through this process. They must have coevolved in complexity with modes of social learning. In this way, the social norms that originally demanded cooperation or punishment could have begun as quite simplistic, such as cooperative hunting, and then coevolved to support more complex behaviours, such as the division of labour. Further studies into the development of social learning of norms, both in the modern era and from an evolutionary perspective, are necessary for a more complete picture of strong reciprocity.

Second, *how do children learn, apply, and calibrate normative behaviour?* Developmental psychology suggests children first acquire norms in contextually specific circumstances and then apply them in different circumstances. Experience then calibrates their understanding of the nature and contextual limits of the norm (Lancy, 1996; Sterelny, in press; Want & Harris, 2002). I suggest that this could be somewhat similar to the conditions of early social learning in the Pleistocene; cooperative and punishment behaviours based on social norms were developed and calibrated at both the individual and group level in a series of trial and error interactions. If this is true, current models that see strong reciprocity as a simply acquired invading behaviour are flawed because they assume the domain of the behaviour is static rather than being constantly expanded and fine-tuned depending on social and environmental context. Social learning then is an important component in the transmission of strong reciprocity. Greater integration of literature in developmental psychology —

including the acquisition and internalization of social norms through teaching, conformism, and usage — is required for a more cogent evolutionary model of strong reciprocity.

Conclusion

Strong reciprocity is a complex of interrelated coevolving traits and needs to be considered as such in evolutionary models. Most work in this field has emerged from experimental economics, a discipline focused on contemporary social institutions and problems. Creating an evolutionary model of human cooperation was not the original intent when these economics games were developed but it is now a promising by-product. Future developments have great potential; including a greater historical focus to explain the evolutionary story as well as further attention on other component parts not discussed above, including but not limited to neural structures and social contexts. Further considerations could also take into account non-human primates; for example, recent research suggests dominant macaques behave pro-socially to emphasize their dominance (Massen et al., 2010). All of these paths can strengthen evolutionary models of strong reciprocity and present a more cogent story of human cooperation.

It is important to note that strong reciprocity is not mutually exclusive from other cooperative behaviours and models. There is nothing to suggest that preference for cooperation with kin can override some tendencies of strong reciprocity or vice versa. In fact, I believe that when theory has developed enough to sustain a complex picture of human cooperation, a multi-level evolutionary story is quite likely. The starting point must be, however, on a smaller-scale; it must begin with papers like this, which conceive the models we already have as complex behaviours, before mapping them all together.

That strong reciprocity is a behaviour has been established. That it is relatively similar around the world has been established. Now questions about origins, its genetic and cultural locus, and the nature of transmission await answers. This paper has sought to guide such enquiries by deconstructing what has so far been conceived as a singular behaviour into a complex suite of interrelated coevolved behaviours. Only with a complex understanding of the nature of human cooperation can we recreate what was surely a complex evolutionary development.

References

- Bolton, G. E., E. Katok, & A. Ockenfels, 2005. 'Cooperation among strangers with limited information about reputation'. *Journal of Public Economics*, 89(8): 1457-1468.
- Bowles, S. & H. Gintis, 2003. 'Prosocial Emotions'. Accessible online at: <<http://www.umass.edu/preferen/gintis/arrowfes.pdf>>.
- Bowles, S. & H. Gintis, 2006. 'The Evolutionary Basis of Collective Action'. Accessible online at: <<http://www.umass.edu/preferen/gintis/Evolutionary%20Basis%20of%20Collective%20Action.pdf>>.
- Boyd, R., & P.J. Richerson, 1985. *Culture and the Evolutionary Process*. Chicago: The University of Chicago Press.
- Boyd, R., & P. J. Richerson, 2009. 'Culture and the evolution of human cooperation'. *Philosophical Transactions of the Royal Society of the Biological Sciences*, 364: 3281-3288.
- Boyd, R., H. Gintis, S. Bowles, P. J. Richerson, 2003. 'The evolution of altruistic punishment'. *Proceedings of the National Academy of Sciences of the United States of America*, 100(6): 3531- 3535.
- Brandt, H., C. Hauert, & K. Sigmund, 2003. 'Punishment and reputation in spatial public goods games'. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1519): 1099-1104.
- Cameron, L. A., 1999. 'Raising the Stakes in the Ultimatum Game: Experimental Evidence From Indonesia'. *Economic Inquiry*, 37(1): 47-59.
- Carpenter, J., S. Bowles, H. Gintis, & S. Hwang, 2009. 'Strong Reciprocity and team production: Theory and evidence'. *Journal of Economic Behaviour & Organization*. 71: 221-232.
- Clutton-Brock, T. H. & G. A. Parker, 1995. 'Punishment in Animal Societies'. *Nature*, 373: 209-217.
- Denant-Boemont, L., D. Masclet, & C. N. Noussair, 2007. 'Punishment, counterpunishment, and sanction enforcement in a social dilemma experiment'. *Economic Theory*, 33: 145-167.
- Dreber, A., et al., 2008. 'Winners don't punish'. *Nature*, 452: 348-351.
- Efferson, C., et al., 2008. 'Conformists and mavericks: the empirics of frequency-dependent cultural transmission'. *Evolution and Human Behaviour*, 29: 56-64.

- Fehr, E. & U. Fischbacher, 2003. 'The nature of human altruism'. *Nature*. 425: 785-792.
- Fehr, E. & U. Fischbacher, 2004. 'Third-party punishment and social norms'. *Evolution and Human Behaviour*. 25: 63-87.
- Fehr, E., & Henrich, J., 2004. Is strong reciprocity a maladaptation? On the evolutionary foundations of human altruism. In: P. Hammerstein (Ed.), *Genetic and Cultural Origins of Cooperation*. Cambridge: The MIT Press.
- Fehr, E., U. Fischbacher, & S. Gächter, 2002. 'Strong reciprocity, human cooperation, and the enforcement of social norms'. *Human Nature*, 13: 1-25.
- Fischbacher, U. & S. Gächter, 2010. 'Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments. *American Economic Review*, 100(1): 541-56.
- Gächter S. & B. Herrmann, 2008. 'Reciprocity, culture, and human cooperation: previous insights and a new cross-cultural experiment'. *Philosophical Transactions of the Royal Society of the Biological Sciences*, 364: 791-806.
- Gächter S., B. Herrmann, & C. Thöni, 2010. 'Culture and cooperation'. *Philosophical Transactions of the Royal Society of the Biological Sciences*, 365: 2651-2661.
- Gardner, A. & S. A. West, 2010. 'Greenbeards'. *Evolution*, 64(1): 25-38.
- Gintis H., 2000a. 'Strong Reciprocity and Human Sociality'. *Journal of Theoretical Biology*. 206: 169-179.
- Gintis, H., 2000b. 'Beyond *Homo economicus*: evidence from experimental economics'. *Ecological Economics*. 35: 311-322.
- Gintis, H., 2004. 'The genetic side of gene-culture coevolution: internalization of norms and prosocial emotions'. *Journal of Economic Behaviour & Organization*, 53:57-67.
- Gintis, H., 2009. 'Behavioural Ethics'. Accessible online at: <<http://www.umass.edu/preferen/gintis/BehaviouralEthics.pdf>>. Accessed on 10 September 2010.
- Gintis, H. 2011. 'Reply to Binmore: Social Norms or Social Preferences?'. Accessible online at: <<http://www.umass.edu/preferen/gintis/ReplayToBinmore.pdf>>. Accessed on 1 November 2011.
- Gintis, H., S. Bowles, R. Boyd, E. Fehr, 2003. 'Explaining altruistic behaviour in humans'. *Evolution and Human Behaviour*. 24: 153-172.

- Gintis, H., J. Henrich, S. Bowles, R. Boyd, & E. Fehr, 2008. 'Strong Reciprocity and the Roots of Human Morality'. *Social Justice Research*, 21: 241-253.
- Gurven, M., A. Zanolini, & E. Schniter, 2008. 'Culture sometimes matters: intra-cultural variation in pro-social behaviour among Tsimane Amerindians'. *Journal of Economic Behaviour and Organization*, 67(3-4): 587-607.
- Harbaugh, W. T., K. Krause, & S. G. Liday, 2002. 'Children's Bargaining Behaviour'. Available online at: <<http://harbaugh.uoregon.edu/Papers/KidBargaining.pdf>>.
- Harlé, K. M. & A. G. Sanfey, 2007. 'Incidental Sadness Biases Social Economic Decisions in the Ultimatum Game'. *Emotion*, 7(4): 876-881.
- Henrich, J., 2000. 'Does culture matter in economic behaviour: Ultimatum game bargaining among the Machiguenga'. *American Economic Review*, 90(4): 973-979.
- Henrich, J. & R. Boyd, 1998. 'The Evolution of Conformist Transmission and the Emergence of Between-Group Differences'. *Evolution and Human Behaviour*, 19: 215-241.
- Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, & H. Gintis (eds.), 2004. *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. Oxford: Oxford University Press.
- Henrich, J., et al., 2005. "'Economic man" in cross-cultural perspective: Behavioural experiments in 15 small-scale societies'. *Behavioural and Brain Sciences*, 28: 795-855.
- Henrich, J., et al., 2006. 'Costly Punishment Across Human Societies'. *Science*, 312: 1767-1771.
- Hermann, B., C. Thöni, & S. Gächter, 2008. 'Antisocial Punishment Across Societies'. *Science*, 319: 1362-1367
- Hopfensitz, A. & E. Reuben, 2009. 'The Importance of Emotions for the Effectiveness of Social Punishment'. *The Economic Journal*, 119: 1534-1559.
- Humphrey, Nicholas, 1997. Varieties of altruism – and the common ground between them. *Social Research*, 64: 199-209.
- Kalish, C.W., 1998. 'Reasons and causes: Children's understanding of conformity to social rules and physical laws'. *Child Development*, 69: 706-720.
- King-Casas, B., et al., 2005. 'Getting to know you: reputation and trust in a two-person economic exchange'. *Science*, 308: 78-83.

- Kreps, D. M. & R. Wilson, 1982. 'Reputation and imperfect information'. *Journal of Economic Theory*, 27(2): 253-279.
- Lancy, David, 1996. *Playing on Mother Ground: Cultural Routines for Children's Development*. London: The Guilford Press.
- Lehmann, L., K. Bargum, & M. Reuter, 2006. 'An evolutionary analysis of the relationship between spite and altruism'. *Journal of Evolutionary Biology*, 19(5): 1507-1516.
- Marlowe, F. W., et al., 2008. 'More 'altruistic' punishment in larger societies'. *Proceedings of the Royal Society of the Biological Sciences*, 275: 587-592.
- Massen, J. J. M., L. M. van den Berg, B. M. Spruijt, & E. H. M., Sterck, 2010. 'Generous Leaders and Selfish Underdogs: Pro-Sociality in Despotic Macaques. *PLoS ONE*, 5(3).
- Milinski, M., D. Semmann, & H. Krambeck, 2002. 'Reputation helps solve the 'tragedy of the commons''. *Nature*, 415: 424-426.
- Nikiforakis, N., 2008. 'Punishment and counter-punishment in public good games: Can we really govern ourselves?'. *Journal of Public Economics*, 92(1-2): 91-112.
- Oosterbeek, H., R. Sloof, & G. Van de Kuilen, 2004. 'Cultural Differences in Ultimatum Game Experiments: Evidence from a Meta-Analysis'. *Experimental Economics*, 7: 171-188.
- Price, M. E., 2008. 'The resurrection of group selection as a theory of human cooperation'. *Social Justice Research*, 21.
- Rakoczy, H., F. Wameken, & M. Tomasello, 2008. 'The sources of normativity: Young children's awareness of the normative structure of games'. *Developmental Psychology*, 44(3): 875-881.
- Rand, D. G., J. J. Armao, M. Nakamaru, H. Ohtsuki, 2010. 'Anti-social punishment can prevent the co-evolution of punishment and cooperation'. *Journal of Theoretical Biology*, 265(4): 624-632.
- Roth, A. E., V. Prasnikar, M. Okuno-Fujiwara, S. Zamir, 1991. 'Bargaining and Market Behaviour in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study'. *The American Economic Review*, 81(5): 1068-1095.
- Sanfey, A. G., J. K. Rilling, J. A. Aronson, L. E., Nystrom, & J. D., Cohen, 2003. 'The neural basis of economic decision-making in the Ultimatum Game'. *Science*, 300: 1755-1758.

- Sterelny, Kim, in press. *The Evolved Apprentice: The Nicod Lectures, 2008*.
- Sutter, M. & M. G. Kocher, 2007. 'Trust and trustworthiness across different age groups'. *Games and Economic Behaviour*, 59: 364-382.
- Sylwester, K. & G. Roberts, 2010. 'Cooperators benefit through reputation-based partner choice in economic games'. *Biology Letters*, 6(5): 659-662.
- Want, S. C., & P. L. Harris, 2002. 'How do children ape? Applying concepts from the study of non-human primates to the development study of 'imitation' in children'. *Developmental Science*, 5(1): 1-14.
- West, S. A. & A. Gardner. 'Altruism, Spite, and Greenbeards'. *Science*, 327: 1341-1344.
- Wu, J., B. Zhanga, Z. Zhoub, Q. He, X. Zheng, R. Cressman, & Y. Tao, 2009. 'Costly punishment does not always increase cooperation'. *Proceedings of the National Academy of Sciences (USA)*, 106(41): 17448-17451.
- Xiao, E. & D. Houser, 2005. 'Emotion expression in human punishment behaviour'. *Proceedings of the National Academy of Sciences (USA)*, 102(20): 7398-7401.