The Evolution of Morality and the End of Economic Man

Geoffrey M. Hodgson

10 March 2013

Published in the *Journal of Evolutionary Economics*

Hertfordshire Business School, University of Hertfordshire, De Havilland Campus, Hatfield AL10 9AB, UK
www.geoffrey-hodgson.info
g.m.hodgson@herts.ac.uk

KEY WORDS: morality, Darwin, evolution, cooperation, altruism

JEL classification: B52, D01, D64

ABSTRACT

1871 saw the publication of two major treatises in economics, with self-seeking economic man at their center. In the same year Darwin published *The Descent of Man*, which emphasized sympathy and cooperation as well as self-interest, and contained a powerful argument that morality has evolved in humans by natural selection. Essentially this stance is supported by modern research. This paper considers the nature of morality and how it has evolved. It reconciles Darwin's notion that a developed morality requires language and deliberation (and is thus unique to humans), with his other view that moral feelings have a long-evolved and biologically-inherited basis. The social role of morality and its difference with altruism is illustrated by an agent-based simulation. The fact that humans combine both moral and selfish dispositions has major implications for the social sciences and obliges us to abandon the pre-eminent notion of selfish economic man. Economic policy must take account of our moral nature.

The Evolution of Morality and the End of Economic Man

Geoffrey M. Hodgson

In two seminal works, William Stanley Jevons (1871) and Carl Menger (1871) placed individual self-interest at the foundation of economics.[1] Three years later, Léon Walras (1874) built neoclassical general equilibrium analysis upon a similar assumption of self-interest. For the next 100 years or more, self-interested "economic man" was the centerpiece of mainstream economic theory. But in the same pivotal year, Darwin (1871) published a contrasting and evolutionary explanation of cooperative solidarity and morality, which took over one hundred years to be confirmed broadly by theoretical and empirical research.

This fact has been largely neglected by evolutionary economists that have strangely hitherto made little use of (or even reference to) Darwinism. While "evolution" is a broad word, and there is no warrant to confine its meaning to Darwinism, modern evolutionary thinking should define its relationship to Darwinism in explicit terms (Hodgson and Knudsen 2010).

There is now an enormous body of empirical research confirming that humans have cooperative as well as self-interested dispositions.[2] But many accounts conflate morality with altruism or cooperation.[3] The argument here follows Darwin in establishing a distinctive and vital additional role for morality.

Morality remains controversial for moral philosophers, notwithstanding a widespread view among them that moral judgments cannot be treated as matters of mere preference or utility maximization. Morality means "doing the right thing." It entails notions of justice that can trump our preferences or interests. Moral judgments are by their nature inescapable, and would apply to all in the same circumstances. They are buttressed by emotional feelings and reasoned argument. Consequently, morality differs fundamentally from matters of mere convenience, convention or conformism. Moral feelings are enhanced by learned cultural norms and rules. Morality is a group phenomenon involving deliberative, emotionally-driven and purportedly inescapable rules that apply to a community.

---

[1] This essay makes use of material from Hodgson (2013), which includes more extensive evidence and discussion. Therein I also contrast my own views on morality with others. For example, in leading accounts by Etzioni (1988) and Sen (1987) there is no attempt to explain how our moral or other dispositions evolved. By contrast, Binmore (1994, 1998) develops a sophisticated evolutionary theory but reduces morality to matters of mere convention. In writing this paper I thank two anonymous referees, David Harper, Mario Rizzo, Xueqi Zhang, participants at the 2011 GROE (Group for Research in Organisational Evolution) Workshop (at Great Offley in Hertfordshire in England), and others for comments on an earlier version of this essay.

[2] See, for example, Güth (1995), Field (2001, 2007), Henrich et al. (2001, 2004), Hammerstein (2003) and Bowles and Gintis (2011).

[3] See, for example, Fehr and Gächter (2002), Bowles et al. (2003), Boyd et al. (2003), Gintis et al. (2005), Kaplow and Shavell (2007) , Bowles and Gintis (2011).

A society or economy cannot function without moral bonds and rules. Our understanding of social institutions and organizations is inadequate unless we appreciate the moral motivations of individuals within them, and how those institutions help to sustain and replicate these moral sentiments.

Accepting the importance of morality does not mean denying that individuals are self-interested. We are selfish, to a major extent. But we are also moral beings, and our ethical feelings and beliefs – well-formed or otherwise – play an ubiquitous role in our interactions with others, even in the modern acquisitive world of business and consumerism.

It is necessary to bring moral motivations back into the picture. Fortunately there is now a flourishing stream of recent evolutionary studies – spanning several disciplines including anthropology, primatology, philosophy, and economics – that can help to fill the gaps. We can begin to understand the role of morality in sustaining institutions and enhancing social cohesion in any society.

In recent years our understanding of the evolution of cooperation has developed enormously. There are many powerful models and simulations exploring mechanisms and conditions under which altruistic or cooperative individuals can become established in a population. What does a focus on morality add to this? How would the models and simulations have to be changed to accommodate ethical norms? By assuming that behavior largely reflects inherited dispositions and neglecting the possibility of countervailing moral deliberations, prominent models of the evolution of altruism pay insufficient attention to why an intelligent and thoughtful individual should bear costs to help others, even if she was genetically and emotionally disposed to do so.[4] Morality helps to explain why. Although the modelers of cooperation often mention morality, they fail to illuminate its crucial role, typically by forcing matters of ethics and preference together into one utilitarian box. This paper includes a simple simulation which gives a different output when moral algorithms are switched on or off, thus illustrating the distinctive contribution of moral systems.

Another reason why morality is important is that when we turn our attention from the past evolution of morality in humans to the consideration of economic and social policies for the present, then a deeper and more detailed understanding of human motivation is required. By treating motivation as all stemming from individual "preferences" – as is typical in much of the literature on the evolution of human cooperation – approaches to policy are more easily diverted into the narrower channels of material or pecuniary incentives, neglecting moral motivations and appeals to ethical values.

The nature of morality is discussed in the following section. The evolution of morality is considered in sections two, three, four and five. The narrative moves from Darwin's sophisticated view, to the modern proposition that moral feelings have a genetic foundation as well as a heavy dependence on culture and deliberation. Section five presents an agent-based model that adds systemic moral processes to a model by Samuel Bowles et al. (2003). Significant differences in output result. Section six concludes the paper.

1. What is morality?

Morality is complex and controversial. In Darwin's (1871, vol. 1, pp. 87-89) account, morality results from a combination of emotional impulses and thoughtful deliberation. He

[4] See the references in the preceding footnote.

argues that although primitive moral feelings have evolved for millions of years among "the progenitors of man" (1871, vol. 1, p. 162), humans alone have a developed sense of morality:

> A moral being is one who is capable of comparing his past and future actions or motives, and of approving or disapproving of them. We have no reason to suppose that any of the lower animals have this capacity … man … alone can with certainty be ranked as a moral being …

For Darwin, morality emerged in humans upon a long-evolved foundation of instinct and impulse. As noted in the following section, Darwin also saw morality as a social phenomenon, involving social relations and shared values.

Much of the recent theoretical work by economists that attempts to explain cooperation in the real world conflates issues of morality with altruism or cooperation under the description of "social" or "other-regarding" preferences (Bowles and Gintis 2011, Camerer and Fehr 2006, Fehr and Fischbacher 2002, Fehr and Camerer 2007). The assumption of "other-regarding" preferences contrasts with the previously-prominent idea that economic man was entirely selfish. But someone with "other-regarding" preferences is still maximizing *her own* utility, and may be regarded as selfish too. To some extent the work of Bruno Frey (1997) improves on this with his distinction between extrinsic and intrinsic motivation. Morality relates more closely to the latter than the former. Viktor Vanberg (2008, p. 608) shows that once we distinguish between preferences over outcomes and preferences over actions then morally-motivated behavior cannot fit into the former category:

> There is … a significant difference between claiming, on the one hand, that agents evaluate outcomes not only in terms of their own narrowly defined interests but also in terms of how they affect the well-being of other persons, and claiming, on the other hand, that agents are motivated to act in accordance with ethical rules or principles of fairness.

The incapacity of the standard utilitarian calculus – or even the language of preferences – to depict adequately the nature of morality becomes even more graphic if we consult the literature in ethical philosophy.

Many controversies divide moral philosophers. The best we can do here is to pick up rather selectively some relevant threads and in some prominent descriptions of the nature of moral judgment. The leading moral philosopher Richard M. Hare (1952) argued that morality was subject to reason and one cannot hold contradictory ethical judgments. He also maintained that any normative judgment was universalizable in the context to which it pertained, in the sense that anyone proclaiming an "ought" in a particular context was committed to prescribing a similar normative judgment for anyone in any relevantly similar situation. As Mackie (1977, p. 33) put it in his classic account, a moral judgment

> is not purely descriptive, certainly not inert, but something that involves a call for action or for the refraining from action, and one that is absolute, not contingent upon any desire or preference or policy or choice, his own or anyone else's.

In his very impressive philosophical account of the *Evolution of Morality*, Richard Joyce (2006, p. 70) argues on the basis of considerations in the philosophical literature that morality has most or all of the following characteristics:

1. Moral judgments express attitudes (such as approval or contempt) and also express beliefs.
2. The emotion of guilt is an important mechanism for regulating moral conduct.
3. Moral judgments transcend the interests or ends of those concerned.

4. Moral judgments imply notions of desert and justice.
5. Moral judgments are inescapable.
6. Moral judgments transcend human conventions.
7. Moral judgments govern interpersonal relations and counter self-regarding individualism.

These characteristics do not establish a *valid* morality; they instead help us to identify what is a *moral judgment*, whether acceptable or otherwise. The argument in this paper relies on descriptive rather than normative ethics: there is no attempt here to identify the "right" morality, but instead to identify the basic nature of a moral claim. Most religions uphold moral claims, but that does not make them all right or just.

Like Darwin, Joyce emphasizes the role of the emotions as well as deliberation. His point (1) establishes that a moral judgment must involve both beliefs and sentiments, and is not reducible to either alone. If an action is impelled *purely* by emotion and sentiment then – as Darwin understood – it cannot amount to moral motivation. Deliberations and beliefs are also vital, but are themselves insufficient because they must be backed by sentiments or emotions: acting morally is more than calculated conformity to moral rules.

Moral judgments may be rationalized in various ways, but they are more than matters of propositional belief or logical syllogism. Defiance of shared moral rules in a group is often met with emotional hostility. Conformity to them may sometimes bring a warm emotional glow. The emotional dimension of moral rules plays an important role in their evolution and their survival, as I shall discuss further below. Guilt (point (2)) is a particularly important emotion that sometimes emerges after breaches of moral rules, and it too plays a part in the evolutionary process.

Joyce's points (3) through (7) reveal the limitations of typical utilitarian approaches. Moral judgments are not simply expressions of an individual's interests, preferences, sentiments or beliefs. They are also claims to universality in their context, which would apply irrespective of the interests, preferences, sentiments or beliefs of those to whom they are supposed to apply.

As both Mackie and Joyce insist, morality surpasses questions of preference. It is a matter of right or wrong, or of duty, of "doing the right thing," irrespective of whether we like it or not. This is part of what makes us human: we are capable of considering moral rules, and understanding that their observance is more than a matter of personal whim or satisfaction. This dimension is missing in much of economics. Moral values are either ignored or subsumed under matters of utility or preference.

Modern society establishes a fundamental difference between moral rules and other (normative) rules. "Murder is wrong" does not carry the same connotations as "splitting infinitives is wrong" or "in Britain one must drive on the left side of the road." Linguistic and traffic rules are matters of convention; they are non-universal. But punishment may still occur when some conventions are breached. Murder is also punishable, but by contrast it is more than a breach of convention.

Threat of punishment or respect for the law are each insufficient to explain the relatively low frequency of murder and other crimes. Most of us abstain from murder not simply because the probability of severe punishment outweighs any expected benefit. Most of us refrain from murder because we believe that it is *morally wrong*; we would desist even if we lived in a country where murder went unpunished.

While there is a difference between morality and mere convention, some conventional rules seem to acquire a moral imperative when they become laws. They sometimes inherit the force of morality from other purportedly universal moral rules, particularly the need to respect others and to obey the law. While conventions may differ from culture to culture, we often conform to them, partly out of mutual respect or legal responsibility. Hence matters of mere convention can acquire some moral force if they become enshrined in law. If so, they do not necessarily become moral issues themselves, but their observance may acquire moral substance by virtue of their legal status. Consequently the moral legitimacy (or otherwise) of the legal system in the eyes of citizens is crucial (Hodgson 2009).

It is a commonplace observation that what may be a moral rule for one culture may not be so for another. But this does not mean that moral rules are reducible to conventions. They become moral rules because many people believe in them as such, and they jointly uphold them as more than matters of convenience, self-interest or convention. The cultural specificity of some moral judgments does not justify a *normative moral relativism*, where one person's morality is deemed as good as any other.[5]

Advocates of normative moral relativism take a further step and uphold that if an act is regarded as morally permissible in a culture then it must be deemed acceptable, even if it is regarded as wrong elsewhere. To accept such a version of moral relativism is to undermine an essential feature of morality itself – that it is absolute and inescapable. Because this feature is denied, such a moral relativist cannot believe in *any* moral judgment in the terms defined above. Hence normative moral relativists are obliged to become moral nihilists or amoralists.

Just as there are unavoidable ontological commitments in any process of enquiry, for humans in societies there are unavoidable moral commitments. And some moral rules are ethically superior to others. Moral relativists respond that there is no way of knowing what the superior moral rules are. But even if there were no way of finding them, this argument is invalid. As with the epistemic error in ontology (the fact that we can never prove that the real world exists does not mean that there is not a real world outside our senses), there is a similar error in ethics (the fact that we may be unable to identify a superior moral code or prove that it is valid does not mean that a morally superior code does not exist). Ignorance of a valid morality does not mean that it is in principle unspecifiable.

Significantly from an evolutionary perspective, studies show a number of common features of moralities across cultures, notwithstanding other important cultural variations.[6] All cultures regard many acts of harm against others as immoral and invest many acts of reciprocity and fairness with moral virtue. All cultures have moral rules concerning required behaviors specific to particular social positions, roles or ranks. Moral codes restraining individual selfishness are also commonplace. As well as sustaining enormous cultural diversity, genetic and cultural co-evolution has ensured that some specific types of pro-social moral rule have endured.

In summary, and in answer to the question that heads this section, a moral judgment involves and expression of attitudes, beliefs and emotions but are also subject to deliberation concerning matters of fairness or justice. In contrast to standard utilitarian approaches, a moral judgment is more than mere convention; it is inescapable and transcends individual

[5] The uncontroversial observation that different cultures have different ethical codes can be described as *descriptive moral relativism*. See Hodgson (2013) for a discussion.

[6] Bok (1978), Roberts (1979), Brown (1991), Schwartz (1994), Haidt and Joseph (2004), Nichols (2004).

preferences or interests.[7] A moral system refers to shared and interactively reinforced moral values in a society or social group. These definitions are incomplete and imprecise, but they are sufficient for our purposes in this paper.

2. Darwin and the evolution of morality

We require evolutionary explanations of the origin and persistence of morality. Darwin's account of this evolution is remarkably resilient, even in the light of modern research. In his *Descent of Man*, Darwin (1871, vol. 1, p. 162) considered dispositions such as "sympathy, fidelity, and courage" that would advantage one tribe again the other in their struggle for existence, and which had been originally "acquired by the progenitors of man." The accent on sympathy is of course reminiscent of Adam Smith in the *Theory of Moral Sentiments*.[8] Among the other qualities he considers, Darwin (1871, vol. 1, p. 162) listed the disposition to obey those in authority rather than follow individually selfish motives:

> Obedience … is of the highest value, for any form of government is better than none. Selfish and contentious people will not cohere, and without coherence nothing can be effected. A tribe possessing the above qualities in a high degree would spread and be victorious over other tribes ...

Obedience to authority is of course a vital mechanism in the establishment of a system of morality in society. Darwin (1871, vol. 1, p. 166) wrote:

> It must not be forgotten that although a high standard of morality gives but a slight or no advantage to each individual man and his children over the other men of the same tribe, yet that an advancement in the standard of morality and an increase in the number of well-endowed men will certainly give an immense advantage to one tribe over another. There can be no doubt that a tribe including many members who, from possessing in a high degree the spirit of patriotism, fidelity, obedience, courage, and sympathy, were always ready to give aid to each other and to sacrifice themselves for the common good, would be victorious over most other tribes; and this would be natural selection.

Darwin proposed that groups containing individuals that devote themselves to the interests of their group will have an advantage in the struggle for survival. Among humans, binding sentiments of sympathy and solidarity are strengthened by a moral code, typically transmitted by instruction and often sanctified by religion (Wilson 2002).

Darwin's evolutionary explanation of moral sentiments relies to some degree on a notion of group selection, where individual traits that benefit the group are assumed to prosper. Darwin did not counter the objection that selfish individuals would be able to free-ride within an

---

[7] Utilitarians will not accept this. But I cannot develop a critique of utilitarianism here. The behavior of any real entity can be made consistent with some utility function (including, more controversially, apparently "inconsistent" behavior). For utilitarians that is enough to clinch the matter. But for the critics, fitting a function to behavioral data does not explain anything, and it overlooks that which specifically makes us human. For fuller discussions and further references see Smart and Williams (1973), Sen (1987), and Hodgson (2013).

[8] In his *Moral Sentiments* and elsewhere Smith emphasized moral motivations and the importance of justice in economic arrangements (Sen 1987, Evensky 2005). Darwin made notes on the *Moral Sentiments* but there is no evidence he read the *Wealth of Nations.* He famously read Malthus's *Essay on the Principle of Population*, but the pages of the copy of the first volume of *Capital* sent to him personally by Marx remained uncut (Hodgson 1993).

altruistic group, and eventually out-breed the unselfish (Williams 1966, Dawkins 1976). Before the theory of group selection was rehabilitated,[9] Darwin's theory of the evolution of morality was regarded as quaint and outmoded. Darwin's account also suffered because of a long-established and popular contrary view that the foundations of human morality are of recent origin, rather than based on "social qualities … acquired by the progenitors of man."

Darwin was ignorant of the mechanisms of inheritance, including genes. We now understand that genetic must be distinguished from cultural group selection. As Joseph Henrich (2004) shows, both types of group selection are possible in principle under specified conditions, but this does not always necessarily mean that they are always strong. Furthermore, genes and culture interact with one another in specific ways. Culture provides part of the environment in which genes are selected, and our genetic endowment influences cultural evolution (Boyd and Richerson 1985, Durham 1991).

Strong arguments support the notion of cultural group selection among humans (Boyd and Richerson 1985, Henrich 2004). The existence of genetic group selection in our species is more problematic. Although it is possible in principle, it depends on the restriction of inter-group migration and the limitation of genetic mixing between groups. But the evidence among primates is that significant group-to-group migration does occur (De Waal 2006, p. 16). There is also inter-group migration in some contemporary hunter-gatherer societies, and consequently lower levels of genetic relatedness within groups (Hill et al. 2011). We lack any clear evidence on the degree of inter-group migration among early humans, but we have no reason to presume that they differed radically from primates in this respect, although several hypotheses have been developed (Bowles and Gintis 2011).

Consequently, on the basis of existing evidence, the genetic foundations of altruistic and moral feelings seem more likely to have evolved first through mechanisms of kin altruism and then reciprocal altruism. William Hamilton (1964) demonstrated that altruistic genetic dispositions can evolve among closely-related individuals and Robert Trivers (1971) showed how altruism could be reinforced in small groups by mechanisms sustaining reciprocation. Our best guess with current knowledge is that altruistic, cooperative and moral feelings then required the further emergence of a culture, so that they could spread through the group and become reinforced by enduring cultural norms. In short, genetic mechanisms established critical masses of altruists in social groups, leading to the spread of cultural norms sustaining cooperation and to the development of systems of morality that further enhanced the fitness of groups. Genes played a role, but also indispensable was culture, particularly through the inculcation of behavioral norms in children by parents (Palmer and Steadman 1997).

In any case it seems certain that very basic moral feelings have a genetic basis and have evolved in family and kin groups. Because reciprocity and cooperation in such circumstances enhances the fitness of the genes, emotional and other dispositions that aided cooperation and family cohesion also had a survival advantage. Moral sentiments thus evolved on a genetic foundation. But they require structured social interaction to become channeled and expressed. Culture developed these sentiments into a transmitted moral code. Hence the evolutionary origins of morality involved the interplay of genetic and cultural factors. Morality thus has both individual and social aspects. In the following two sections the genetic foundation of morality is discussed in more depth.

---

[9] Major contributions to this rehabilitation include Wade (1978), Wilson (1980, 1983), Wilson and Sober (1994), Sober and Wilson (1998), Wilson and Wilson (2007).

3. Contra Huxley and Dawkins: morality is more than skin-deep

Is morality something that emerged only after the development of complex societies and civilized behavior? Is it a recent innovation that has to be instilled into everyone against their rude and selfish evolved nature? Or is it more than skin-deep?

Darwin insisted that morality proper relies on levels of conscious deliberation, which are developed among humans alone. But at the same time, Darwin saw the biological roots of morality as a product of natural selection. Because of our unique capacities for linguistic communication, humans were the first species to develop and articulate moral codes; but the foundations of morality go far back into our pre-human past.

But evolution and morality have often been regarded as separate issues, even by Darwin's disciples. In 1893 Darwin's friend Thomas Henry Huxley famously argued that "the ethical progress of society" depends on neither ignoring, accepting nor imitating natural selection in that sphere "but in combating it" (Huxley and Huxley 1947, p. 82).[10]

Redolent of Huxley, Richard Dawkins (1976, pp. 2-4, 215) describes the "ruthless selfishness" of our genes. He also claims that "we are born selfish" and "anything that has evolved by natural selection should be selfish." Not only genes but also individuals are described as selfish.[11] Similar claims that evolution gives rise to the self-interested human individual can be found in works by Michael Ghiselin (1974) and Richard Alexander (1987). But Dawkins adds in a contradictory and Huxleyan manner: "Let us try to teach generosity and altruism. … We, alone on earth, can rebel against the tyranny of the selfish replicators." We are born selfish but somehow we can choose to be otherwise. The contrast with Darwin's (1871) views on human nature and morality are graphic.

There are many problems with the Huxley-Dawkins separation of our evolutionary legacy from our current choices and capacities (Midgley 2003, De Waal 2006, Joyce 2006). Above all, if natural selection provides us with selfish dispositions, then why should we be inclined to "combat" selfishness or "teach generosity and altruism," unless they are some sort of elaborate selfish ruse to get the upper hand? Why should nature-red-in-tooth-and-claw rule in one sphere but not another? Furthermore, any inclinations "to teach generosity and altruism" and to "rebel" against our own selfishness are unexplained. We are asked to overturn natural selection in the human domain without the assistance of any evolved human attribute. To "combat" natural selection would also mean suppressing the evolved moral sense and the only possible moral ally in this rebellion against selfishness.

Frans De Waal (2006, p. 10) criticizes the Huxley-Dawkins view where "morality is presented as a thin crust underneath which boil antisocial, amoral and egotistic passions." He describes this as the "veneer theory" of morality and presents contrary evidence from his studies of primates that they are capable of sympathetic and cooperative emotions. Hence our

---

[10] Huxley's severance of the evolution of ethics from the evolution of humanity was countered by a number of writers including Westermarck (1891) and Kropotkin (1902).

[11] This is a major inconsistency in Dawkins's (1976) position. The "selfish gene" metaphor supports his notion that evolutionary selection prioritizes successful genes because these are the key bits of information that replicate and program the growth and behavior of organisms. Whatever the merits and demerits of this "genes-eye view" of evolution, it cannot sustain the notion that individuals are entirely selfish as well. Inconsistently with his selfishness claim, Dawkins cites and endorses key arguments by Hamilton (1964) and Trivers (1971) that show that the individual may bear costs or even sacrifice itself, as if it were acting instead for the sake of its genes.

moral capacities are grounded in our evolution as a social species, over many millions of years.

On one side is the "veneer theory" view that evolution leads to individual selfishness only, and morality is a desirable but optional overlay. On the other side, the critics argue that our long survival as a social species has meant the evolution of social instincts such as fairness and sympathy and the emotional capacity for guilt, and these form part of the foundations for morality. The critics cite evidence from the study of primate behavior, from psychology, from neuroscience, and from experimental economics.[12]

The evidence of the critics is compelling. But we may not conclude that a fully-developed morality has existed for millions of years. Morality proper requires the capacity to compare and deliberate upon actions and motives. It requires communication with others and a highly developed language (Joyce 2006). As Darwin (1871, vol. 1. p. 72) wrote: "after the power of language had been acquired and the wishes of the same community could be distinctly expressed, the common opinion how each member ought to act for the public good, would naturally become to a large extent the guide to action."

Morality proper, involving articulated abstract rules, could not have developed earlier than language. A sufficiently well-developed language is of fairly recent origin – the earliest estimates being hundreds of thousands of years (Oppenheimer 2004) and the more recent being around 50,000 years ago (Diamond 1991). By contrast, the genus homo has been in existence for several million years. So is morality a recently-acquired "veneer" after all? To answer this question we need to distinguish between, on the one hand, the evolution of the preconditions for the development of moral systems and, on the other hand, the full emergence of morality itself. As outlined in the next section, some of the preconditions of our morality evolved as early as our ape-like ancestors.

4. More on how morality evolved

The nuts and bolts of morality include capacities for sympathy and cooperation that we share with the primates. The evidence suggests that much of this is biologically inherited as social instincts. But long before the evolution of human language, a form of proto-cultural and non-genetic transmission occurs. De Waal (1996, 2006) argues that primates can read or even transmit emotional states such as approval, empathy and fear, through sounds, body language, facial expressions, and pheromonal excretions. This is known as "emotional contagion." The communication of emotional states is at the core of the capacity for empathy: apes can understand and even share the joys or sufferings of others.[13]

The possibility of reading emotions in others, and even replicating emotional states, is a crucial inheritance mechanism for transmitting useful information and enhancing social cohesion. For example, the transmission of fear in the group can lead to collective flight from a predator. Such mechanisms are ubiquitous in nature. The capacities to respond to such signals are genetically inherited as instincts. Organisms are genetically programmed to respond to signals that are relevant for their survival.

---

[12] See, for example, Güth (1995), De Waal (1996), Greene and Haidt (2002), Zak (2004), Tancredi (2005), Hauser (2006), Fehr and Camerer (2007).

[13] Consider also *neural mirroring*, where emotional contagion is shown to be a result of "mirror neurons" in the brain (Rizzolatti et al. 2007).

Among the repertoires of response for sophisticated organisms is the imitation of the behavior of others. There are at least two types of imitation among humans and primates. The first is conformist transmission (Boyd and Richerson 1985). It has been shown that genes disposing individuals to such conformism would be selected in some contexts (Henrich and Boyd 1998). A second psychological mechanism is prestige-based imitation (Henrich and Gil-White 2001, Henrich 2004), where individuals learn advantageously from the more successful. This second mechanism must involve capabilities to recognize social hierarchy and prestige. In any social species such instinctive propensities are likely to be selected over time; they would bestow survival advantages for the individual and the group.

There is also evidence for learned or inherited dispositions, which are triggered in specific contexts, to punish those who break the rules or fail to enforce them.[14] The relevant inherited dispositions have evolved in our social species over millions of years. Some punishment involves "strong reciprocity" (Gintis 2000) where there is a propensity not only to punish cheats, free-riders, rule-breakers and self-aggrandizers, but also to punish others who fail to punish the offenders. Especially within small groups, these propensities are driven by strong emotional feelings of anger. There is also evidence of such dispositions among primates (De Waal 1996). In a complex culture, emotionally-empowered rules can help to enhance notions of justice and morality (Darwin 1871, De Waal 2006, Robinson *et al*. 2007). Given these emotions were present in our ape-like ancestors, morality has a genetic as well as a cultural foundation. Genes are insufficient to generate a moral system, but the cultural phenomenon of morality is fuelled by biologically-grounded emotions and value impulses.

Support for this view comes from experiments showing that preverbal babies assess individuals in the light of their behavior towards others. Studies of six- and ten-month-old infants show that they prefer an individual who helps over one who hinders another, prefer a helping individual to a neutral individual, and prefer a neutral to a hindering individual (Hamlin et al. 2007). Such inclinations may serve as a foundation for later moral thought and action, and their early developmental emergence supports the view that such capacities have an inherited biological grounding. Other studies show that infants as young as four years have a sense of fairness (McCrink et al. 2010). Young children are capable of distinguishing moral from prudential norms (Tisak and Turiel 1984), although many adults would have difficulty articulating the distinction.

Moral development starts in very young children and is subject to abrupt developmental leaps. For example, ideas concerning fairness tend to emerge around the age of four years, to be applied wildly to multiple contexts (Fiske 1991). As Jonathan Haidt (2001, p. 827) concludes: "This pattern of sudden similarly timed emergence with overgeneralization suggest the maturation of an endogenous ability rather than the learning of a set of cultural norms."

Through our genes we inherit the capacity to quickly respond to social dilemmas by developing emotions. These emotions not only dispose us to make choices but help us to form rapid judgments concerning what is morally right or wrong. Moral judgments help us justify our actions to others and to exhort others to approve or imitate. Genetic dispositions to deal

[14] See Boyd and Richerson (1992), Andreoni (1995), De Waal (1996), Ben-Ner and Putterman (2000), Fehr and Gächter (2000a, 2000b, 2002), Gintis (2000), Field (2001), Price et al.(2002), Boyd et al. (2003), Carpenter et al. (2004), Gintis et al. (2005), Wiessner (2005), Henrich et al. (2006), Fehr and Gintis (2007), Guzmán et al. (2007), Carpenter and Matthews (2009) , Henrich et al. (2010), Bowles and Gintis (2011).

with social dilemmas by developing emotionally-charged value intuitions can thus have strong survival value.

Our genes do not tell us what is moral or immoral. We try to learn that through engagement in a social culture. The foundations of our moral capacity have evolved over the millions of years that we have been a social species. But in the last hundred thousand years or so, the full development of human moral capacities has been highly dependent on particular cultural settings, allowing for multiple and contrasting moral systems on the basis of a common instinctive bedrock.

Given that we also inherit genetically a long-evolved capacity to imitate others and a capacity for empathy, we are likely to conform to strong moral claims, especially when made by high-status or numerous individuals. Cultural mechanisms lead to conformity within the group. Morality thus operates in a group context. Although rebellion against the prevailing moral rules is possible, clashes are more likely between groups with different cultures.

The evolution of moral capacities is much like the evolution of language, as well as depending on language itself. Because it would be impossible for an infant to learn all the rules of a language solely by reinforcement learning and cultural transmission, some very basic linguistic capacities must be inherited as instincts.[15] Language acquisition depends on both inherited genetic capacities and interaction with others in specific cultural contexts. Similarly, the evolution of morality is "half-art and half-instinct." When we judge an action as morally right or wrong, in part we rely on instincts and feelings. But this does mean that evolution creates a single morality for humankind. Variation between communities in their expressed moral norms is like variation between cultures in their spoken languages (Hauser 2006, pp. 419-20).

In sum, morality depends on language, communication of abstract concepts, discussion and reflection upon them, and the derivation of principles that are general (at least for the group). We inherit some core moral capacities genetically, but not morality as a whole. The development of morality proper depends on interaction with others in a social context.

5. Modeling the evolution of moral systems

The development of moral dispositions in individuals is necessarily a matter of both inherited value-intuitions and developmental processes within a social culture and community. Morality is more than an individual attribute; hence some moral philosophers refer to *moral systems* (Thiroux and Krasemann 2009). These systems include mechanisms of moral inculcation and enforcement in the group. Morality depends on social relations.[16] Hence, it cannot be understood *simply* as individual attributes or dispositions, however complex or multidimensional. It cannot be encompassed adequately by the "social" or "other-regarding" preferences of an individual, or by a "multiple self" with plural preference functions (Margolis 1982, Elster 1986). While individuals and their attributes must be part of the

[15] But their extent is a matter of dispute among psychologists and linguists (Pinker 1994, Deacon 1997, Sampson 2005, Evans and Levinson 2009).

[16] As Arrow (1994) argues forcefully, while economists often make contrary claims, all known economic analyses involve social relations or structures as well as individuals. Hodgson (2007, 2013) shows that "methodological individualism" is a highly ambiguous and lamentably imprecise doctrine, and plausible versions treat social relations as well as individuals as part of the *explanantia*.

explanation, we also have to bring in the role of structured relations between individuals (Hodgson 2007). Morality is both a biological and a social phenomenon, and is irreducible to either individual preferences or genetic endowments.

Moral systems depend on social positions and relations of authority. Morality is sustained not merely through genetic and cultural transmission, but through the replication of structures of authority and power. According to Haidt and Joseph (2008), respect for authority is one of our five evolved value-intuitions. It has coevolved with these structures, so together they provide powerful systems of social cohesion, as well as sources of inequality and oppression. Hence what is involved is not only imitation or learning, but also the reproduction of institutions and cultural mechanisms for legitimating positions of authority (Runciman 2002).

The evolution and sustenance of moral systems involves a battle between the selfish inclinations of individuals and group moral pressure to conform and cooperate. If the group pressure is insufficient, then individual fitness will be enhanced by acting selfishly. If the group moral pressure restrains selfishness sufficiently, then individual fitness will be enhanced by cooperating with others. It is a classic case of individual fitness being context dependent. In such circumstances we can also speak of group fitness.

To what extent is it possible to model the evolution of moral communities? Because morality depends on both biologically inherited feelings and cultural transmission, we are reminded of the models of *dual inheritance* pioneered by Robert Boyd and Peter Richerson (1985) and others. Samuel Bowles and Herbert Gintis (2011) have also developed useful models concerning the evolution of cooperative and altruistic behavior. Bowles and Gintis (2011, p. 125) point out that when evolutionary processes involve multiple and complex interacting levels, then analytically soluble models are out of reach. So they resort to agent-based simulations. Bowles, Jung-Kyoo Choi and Astrid Hopfensitz (2003) simulate the coevolution of altruism and cooperation-enhancing institutions in a population. Such work needs to be improved by including moral motivation.

Morality fills a serious gap in most models or simulations of the evolution of altruism. Typically it is assumed that there two types of person: altruistic cooperators versus self-interested egoists. Initially there is a low or zero frequency of altruists, and more come into existence by replication or low-probability random mutation. Altruists help others at their own cost. But even with low initial frequencies and relatively high costs, these models and simulations show that a population with a much higher frequency of altruists can evolve.

Apart from the low-probability random mutation, individuals in these models are assumed to remain as one type or the other. But what these models do not show is why any intelligent and reflective individual should continue to act in accordance with any inherited disposition towards altruism or cooperation. Why should altruists remain altruists, or cooperators remain cooperators? The models assume that we inherit a fixed inclination, and there is no role for countervailing deliberation.

Most early models of the evolution of altruism were developed to apply to non-human species, with examples such as the alarm call of an individual bird that warns the flock of the approach of a predator, thereby placing itself at greater risk. These models typically assume a close correlation between genetic dispositions and actual behavior. But in the case of humans, this is much less tenable. Culture plays a more important role. And humans have a much greater capacity to reflect on their circumstances and consider the consequences of their actions. We are neither cultural nor biological puppets. We can reflect upon our situation and resist our emotions.

This exposes a major limitation of previous models. They fail to take into account that emotional feelings such as sympathy or guilt can be countered by deliberation upon our self-interest. If an individual experiences (genetically or culturally acquired) feelings of sympathy for an orphaned child or a destitute beggar, or anticipates the warm inner glow of satisfaction from the envisaged act of giving, then they can reflect upon their costly sacrifice and override these emotions. We are deliberative and thoughtful, so we do not always follow our emotions. We do not always answer the call of sexual lust by making advances on an attractive partner, so why shouldn't we also resist the pangs of sympathy and maintain our bank balance? We have basic urges, but we often resist them. So if we can repel our desires, then why cannot we resist the altruistic urge as well? Some additional factor is required to explain the enduring frequency of altruistic or cooperative behavior. This factor is typically absent from the models, or it is subsumed under individual altruistic dispositions. Morality is this missing ingredient.

A system of morality works to some extent by dovetailing with altruistic emotions. Insofar as these emotions have a genetic foundation they will reflect communalities of human evolution and sustain some ethical universals. Other emotions will reflect specific cultural settings. A disposition to respond to all such value-laden emotions by reflecting in terms of "doing the right thing" or "do unto others as you would have them do unto you" would help to sustain altruistic sentiments and counter selfish thoughts about personal costs and benefits.

Groups with moral systems that suitably enhance cohesion and cooperation can have fitness advantage over other groups with less effective moral systems. This would be a process of cultural group selection. It would be the selection of groups as objects in a population, with selection outcomes favoring particular customs and habits that sustained fitness-enhancing ethical principles. At the same time this cultural process depends on the prior and sustained evolution of particular genetic dispositions and capacities. These evolved initially as a result of inclusive fitness, and they were then enhanced by pro-cultural developments in our early ancestors.

Aspects of morality can be illustrated by a simple simulation. The strategy is first to replicate the model of the evolution of altruism developed by Bowles, Choi and Hopfensitz (2003) – henceforth BCH-2003 – but with some minor modifications noted below. A switch is added to this model that enables some moral interactions, outlined below. It is shown that they make a major difference to the outcomes of the simulations.

The BCH-2003 model is meant to illustrate the evolution of cooperation and altruistic dispositions among humans in the last few hundred thousand years. The model assumes a population of agents consisting of lifetime altruists and non-altruists. They are divided into 20 groups. For each simulation, total population size is given and group size is approximately constant, modified only by random migration among groups and by the outcomes of conflict between groups. Each individual also has access to baseline material resources resulting from equal access to further resources, such by as hunting or foraging.

Each group has evolving institutional characteristics, termed segmentation (s – the tendency to interact with similar individuals in the group rather than others) and taxation (t – the proportion of revenue taxed from individuals and then shared equally among the group). Each simulation starts with altruists, taxation and segmentation set to zero. Subsequently, for each group, t and s are altered (up or down) each generation by a small random mutation.

In each period in the BCH-2003 model, each member of a group is randomly paired within the group to play once a Prisoner's Dilemma game. The member pairs with a similar type with probability s and is paired with dissimilar types with probability $1 - s$. The resulting

fitness of every agent is the sum of its game payoff plus 10 baseline units. Replicas of the current generation (with possible mutations) constitute the next generation. They are produced by drawing (with replacement) from the current group membership with the probability that any member will be drawn equal to that member's share of the total payoffs of the group. With a low probability 0.001 a member of the next generation is not a replica of its parent, but is an altruist or non-altruist with equal probability. Then there is inter-group migration: with probability 0.2 each member of the new generation relocates to a group randomly selected from the other groups.

A high frequency of altruists contributes to a group's success in intergroup conflicts, allowing it to increase in size. An individual's expected share of the group's next generation's offspring is equal to the individual's share of the group's total payoffs. Because non-altruists do better in within-group interactions they produce more offspring.

With probability 0.25 each group is selected and among those selected lethal competition takes place between randomly paired groups. The group with the higher total payoff wins. Members of the losing group all die and the winning group populates the site occupied by the losers with replicas of its members. The winning group splits with members assigned randomly into two new groups. The new inhabitants adopt the institutions of the group from which they descended.

Institutions are also subject to stochastic variation, increasing or lowering *t* and *s* each period. Both segmentation and resource sharing impose costs on the groups adopting them. More segmented groups may fail to capture the benefits of diversity or of economies of scale, and resource sharing may reduce incentives to acquire the resources to be shared.

Figure 1 shows the results from a modification of the BCH-2003 model that incorporates both altruism and morality. The model here is very similar to BCH-2003, except instead of being altruists or non-altruists, agents have variable altruistic dispositions on a scale from zero to unity. Furthermore, this model explicitly adds the effects of morality as a social system, which is absent or subsumed under individual altruism in BCH-2003. As in their model, agents are place randomly in several groups. The simulation goes through several generational cycles. At the start of the simulation, the levels of individual altruism, individual moral sensitivity, group segmentation, group taxation and group morality are set at zero for all individuals and groups.

Consider the institutional variables. As in BCH-2003, *t* and *s* are altered (up or down) each generation by a small random mutation. But in the present model each group's morality level is also altered (up or down) by a small random mutation positively adjusted by the average levels of morality and altruism in the group. It is thus a bottom-up process.

Moral sensitivity is the tendency at the individual level to take group morality levels into account when playing the game. An enduringly virtuous person in a selfish crowd, and a stubbornly selfish person in a group of altruists, are thus both morally insensitive.

Crucially, morality is introduced in this modified model by the combination of group morality and (variable) individual moral sensitivity, and by their interactions with altruism. This is a small but significant alteration, and illustrates the systemic role of morality in an extremely simple manner.

In this present model, at the beginning of each generational cycle, each individual's altruistic disposition is increased by a small, randomized increment that is proportional to the group level of morality. (The previous altruistic disposition is multiplied by the sum of unity plus a randomized positive fraction – chosen with equal probability in a small interval – of

group level morality, with results above unity truncated to unity.) This top-down process represents institutional processes of individual moral instruction in groups with some positive level of morality. Each individual's moral sensitivity is raised or lowered by a small randomized increment depending on whether its value is respectively below or above the group level of morality. (The previous moral sensitivity is adjusted by the difference between individual moral sensitivity and group morality, multiplied by a randomized fraction chosen with equal probability in a small interval, with results above unity or below zero truncated to unity or zero respectively.) Upward movements in this case correspond to conformist socialization or instruction. As with the BCH-2003 model, all institutions (including group morality in this modified model) are maintained at a cost to group fitness.

As with BCH-2003, all individuals within each group are paired randomly to play a Prisoner's Dilemma game. Also here, the group's segmentation value regulates the probability that two identical agents will be paired. But in the present model two similar and non-identical argents also have a chance of pairing, with higher segmentation making the similarity conditions more stringent than they are with lower segmentation. In other words, the chances of two individuals being paired that have a given degree of similarity to one another (in terms of their levels of altruism and moral sensitivity) are decreased by greater overall segmentation. And given the degree of overall segmentation, the chances of any two individuals being paired increase with their similarity.

The chances of cooperating in the game are related to the individual's altruistic disposition, enhanced by a term involving the product of individual moral sensitivity and the (institutional) level of morality in the group as a whole. A small random component also affects outcomes. Although thoughtful moral reflection is not itself modeled, this combination of individual dispositions, group influences and additional variation is crudely consistent with a role for deliberation. This is excluded by the BCH-2003 model, where born (non-)altruists always behave as (non-)altruists.

The game is then played and the payoffs determined. As in BCH-2003, it is assumed that the resulting fitness of every agent is the sum of its game payoff plus 10 baseline units. The next steps concerning individual replication, generational replacement, possible migration, inter-group combat, group elimination, group replacement, and group division are also very similar to the BCH-2003 model.

The most significant differences between the present and the BCH-2003 model are the replacement of fixed individual binary altruistic dispositions by those that are continuously variable over a 0-1 interval, the replacement of complete by partial influence of altruistic dispositions on behavior, the top-down effects of group morality levels on individual altruism and individual moral sensitivity, and the bottom-up effect of average individual altruism and average moral sensitivity on group morality.

In some circumstances, this combination of top-down and bottom-up feedback effects can lead to virtuous circles of moral development, involving ongoing positive feedbacks. Emergent clusters of cooperating altruists will include those with high levels of altruism plus some with relatively high levels of moral sensitivity who became more altruistic as a result of contact with other altruists. These clusters can positively influence the level of morality in the group. In turn this helps to raise the level of altruism for other members of the group. A process of virtuous positive feedback can sustain and reinforce altruism in the group as a whole. This morality-based positive feedback is absent in BCH-2003.

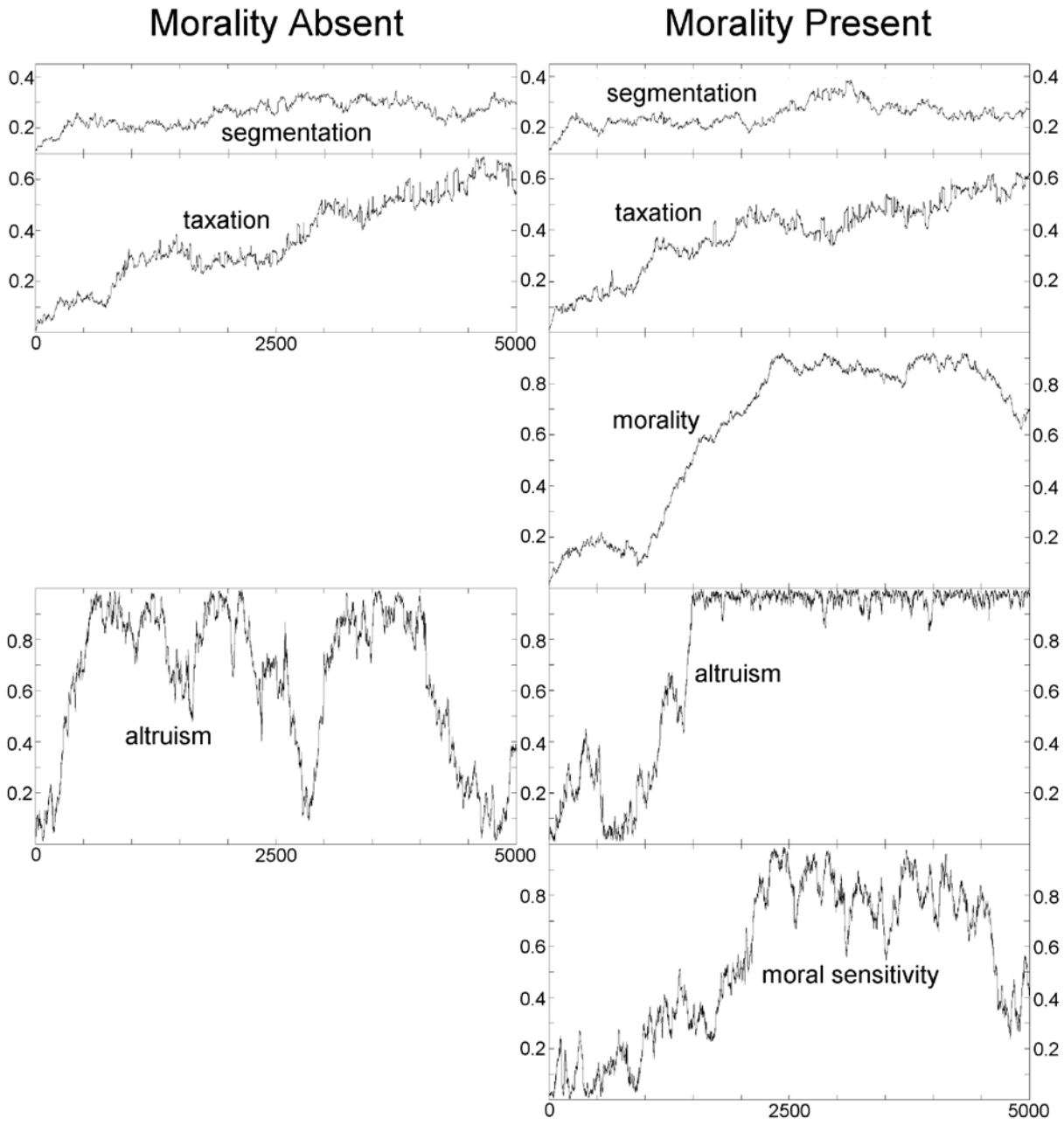

FIGURE 1. The Evolution of Altruism, Morality and Institutions

The figure shows the results of two runs of an agent-based model, with 200 replicating heterogeneous agents and 10 varied groups, each run being 5000 generations (horizontal axes). Panels for segmentation (the tendency to interact with similar individuals in the group), taxation (group sharing) and morality show the average of these group institutional variables. Panels for altruism (the basic tendency to cooperate in a Prisoner's Dilemma game) and moral sensitivity show the averages through time for these individual attributes in the whole population. The three panels on the left show outputs for a run when moral institutions were absent, as in the Bowles et al. (2003) model. The five panels on the right show outputs for a simulation with the morality algorithms in the program are switched on, and the moral variables are allowed to change and take effect. The same model is outlined in Hodgson (2013) but for comparative purposes the simulation runs shown here are different.

Examine Figure 1 and compare the results for the simulations without and with moral systems in place. As in BCH-2003, the three panels on the left show cycles of boom and bust for altruism and other variables. Altruism reaches high levels for several generations, and then crashes again.[17] If most people are altruistically disposed, then maverick non-altruists can exploit the situation and reap higher rewards. These free-riders will reproduce more rapidly and bring the average level of altruism crashing down. By contrast, the panels on the right-hand side of the figure show that when the moral system is in operation, then a high level of average altruism, once achieved, can be sustained by moral imperatives. In these circumstances, individual moral sensitivity and group taxation also reach higher levels. But given that taxation – like segmentation and morality – is a costly institution for the group to maintain, there are burdens on its rise.

To obtain a more robust comparison of runs with and without the morality component, 40 runs in each mode were averaged. The results (not reproduced here) show that mean altruism reaches significantly higher levels with the moral system operative. Without morality, after 1300 generations the mean altruism level stabilizes at about 0.5. But when morality is operative, after 2300 generations mean altruism remains above 0.9. When operative, the global morality level on average rises and stays above 0.8 after about 2500 generations. Consequently the patterns in Figure 1 are quite typical of many runs with the same parameter values.

It is not claimed that the simulation adequately encapsulates morality and moral systems. It is far too simple. Deliberative processes are crudely represented at best. But sufficient features of a moral system are introduced to show major differences with the previous state of the modeling art.

All these runs are illustrative and do not reproduce anything like the exact course of human evolution. The evolution of moral systems in this model is also sensitive to parameter values. The point of the exercise is not to predict actual outcomes but to indicate how the evolution of moral systems may begin to be modeled and to illustrate the difference between moral factors an individual altruism. Morality has to be considered as a group and relational, rather than simply an individual phenomenon.

6. The bottom line: economics without economic man

Morality, as understood by leading moral philosophers, cannot be incorporated into models based on individual utility maximization, even with “social” or “other regarding” preferences. This is because morality is about “doing the right thing,” even if it would otherwise not be the preferred option. Moral judgments are inescapable and cannot be reduced to mere preferences. I argue elsewhere that any form of behavior can be fitted into a utility function. But fitting behavior into functions does not explain its cause or motivation.[18] Digging deeper into the evolutionary and cultural origins of our motives leads us to the issue of morality. Moral motives have evolutionary origins and are sustained through interaction with others: morality is a social as well as an individual phenomenon. This crucial point has been

[17] Compare with the figures in Bowles et al. (2003, pp. 142-3).

[18] See Hodgson (2013, esp. ch.3). Even apparently inconsistent or intransitive preference rankings can be forced into a utility function, by taking into account that different choices always take place in an (at least slightly) different context or with different information.

demonstrated with a simple agent-based simulation model, where the moral algorithms can be turned on or off for comparison.

Humans have moral capacities. But we are also self-interested. Evolution has provided us with instincts that trigger our moral development in suitable socio-cultural settings, and with basic instincts such as hunger and lust that can be spurs to egoism. Through our socialization we typically develop into complex personalities where all biologically inherited impulses are extended or constrained to different degrees and in different ways.

The diverse inner impulses that we bring into the world may come into conflict as our personality develops, in the institutional settings of parental care, peer group interaction and organized education. These settings have major effects on how the moral and self-interested aspects of our personalities develop. Given our declining potential for adaptation as we get older, the earliest years are the most formative.

While accepting that individuals have multi-faceted personalities, the mistake of several influential neoclassical economists was to assume that in the economic sphere self-interest was overwhelming, and our altruistic and moral tendencies could be ignored as we entered the world of contract and business (Jevons 1871, Edgeworth 1881, Wicksteed 1910). Writers such as Gary Becker (1976) claimed that utility maximization, developed in the neoclassical analysis of business life, applies generally to all social interactions.

But even firms and markets are unavoidably infused with moral considerations (Hobson 1929, McCloskey 2006, Zak 2008). These may be countered or developed by example or circumstance. If policy-makers ignore our moral dispositions and concentrate on self-interest alone, then they will threaten the very fabric of a modern market economy (Schumpeter 1942). The acknowledgement of moral motivation opens a large agenda for economists. It is highly relevant for the theory of the firm (Minkler 2008, Lopes et al. 2009, Hodgson 2013). The argument in this paper shows that morality cannot be reduced to individual preferences or altruism. Economic policy is not just about maximizing satisfaction while ensuring that no-one's utility is decreased; it should be about guiding and enhancing our moral dispositions. Especially from an evolutionary perspective, and even in the competitive world of modern business, there is no excuse for ignoring the evolution of moral systems and the moral motivations of economic agents.

References

Alexander, Richard D. (1987) *The Biology of Moral Systems* (New York: Aldine de Gruter)

Andreoni, James (1995) 'Cooperation in Public Goods Experiments: Kindness or Confusion?' *American Economic Review*, **85**(4), December, pp. 891-904.

Arrow, Kenneth J. (1994) 'Methodological Individualism and Social Knowledge', *American Economic Review (Papers and Proceedings)*, **84**(2), May, pp. 1-9.

Becker, Gary S. (1976) *The Economic Approach to Human Behavior* (Chicago: University of Chicago Press).

Ben-Ner, Avner and Putterman, Louis (2000) 'Some Implications of Evolutionary Psychology for the Study of Preferences and Institutions', *Journal of Economic Behavior and Organization*, **43**(1), September, pp. 91-9.

Binmore, Kenneth (1994) *Playing Fair: Game Theory and the Social Contract. Volume I* (Cambridge, MA: MIT Press).

Binmore, Kenneth (1998) *Just Playing: Game Theory and the Social Contract. Volume 2* (Cambridge, MA: MIT Press).

Bok, Sissela (1978) *Lying: Moral Choice in Public and Private Life* (New York: Pantheon).

Bowles, Samuel, Choi, Jung-Kyoo and Hopfensitz, Astrid (2003) 'The Co-Evolution of Individual Behaviors and Social Institutions', *Journal of Theoretical Biology*, **223**(2), July, pp. 135-47.

Bowles, Samuel and Gintis, Herbert (2011) *A Cooperative Species: Human Reciprocity and its Evolution* (Princeton, NJ: Princeton University Press).

Boyd, Robert, Gintis, Herbert, Bowles, Samuel and Richerson, Peter J. (2003) 'Evolution of Altruistic Punishment', *Proceedings of the National Academy of Sciences*, **100**(6), March, pp. 3531-35.

Boyd, Robert and Richerson, Peter J. (1985) *Culture and the Evolutionary Process* (Chicago: University of Chicago Press).

Boyd, Robert and Richerson, Peter J. (1992) 'Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups', *Ethology and Sociobiology*, **13**, pp. 171-95.

Brown, Donald E. (1991) *Human Universals* (New York: McGraw-Hill).

Camerer, Colin F. and Fehr, Ernst (2006) 'When Does "Economic Man" Dominate Social Behavior?', *Science*, **311**, 6 January, pp. 47-52.

Carpenter, Jeffrey P., Matthews, Peter Hans and Ong'ong'a, Okomboli (2004) 'Why Punish? Social Reciprocity and the Enforcement of Prosocial Norms', *Journal of Evolutionary Economics*, **14**(4), October, pp. 407-29.

Carpenter, Jeffrey P. and Matthews, Peter Hans (2009) 'What Norms Trigger Punishment?' *Experimental Economics*, 12(3), September, pp. 272-88.

Darwin, Charles R. (1871) *The Descent of Man, and Selection in Relation to Sex*, 2 vols (London: Murray and New York: Hill).

Dawkins, Richard (1976) *The Selfish Gene* (Oxford: Oxford University Press).

De Waal, Frans B. M. (1996) *Good Natured: The Origin of Right and Wrong in Humans and Other Animals* (Cambridge, MA: Harvard University Press).

De Waal, Frans B. M. (2006) *Primates and Philosophers: How Morality Evolved* (Princeton: Princeton University Press).

Deacon, Terrence W. (1997) *The Symbolic Species: The Co-Evolution of Language and the Brain* (New York: Norton).

Diamond, Jared (1991) *The Rise and Fall of the Third Chimpanzee* (London: Vintage).

Durham, William H. (1991) *Coevolution: Genes, Culture, and Human Diversity* (Stanford: Stanford University Press).

Edgeworth, Francis Y. (1881) *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences* (London: Kegan Paul).

Etzioni, Amitai (1988) *The Moral Dimension: Toward a New Economics* (New York: Free Press).

Evensky, Jerry (2005) 'Adam Smith's Theory of Moral Sentiments: On Morals and Why They Matter to a Liberal Society of Free People and Free Markets', *Journal of Economic Perspectives*, **19**(3), Summer, pp. 109-30.

Elster, Jon (1986) *The Multiple Self* (Cambridge: Cambridge University Press).

Evans, Nicholas and Levinson, Stephen C. (2009) 'The Myth of Language Universals: Language Diversity and its Importance for Cognitive Science', *Behavioral and Brain Sciences*, **32**(5), pp. 429-92.

Fehr, Ernst and Camerer, Colin F. (2007) 'Social Neuroeconomics: The Neural Circuitry of Social Preferences', *Trends in Cognitive Science*, **11**(1), pp. 419-27.

Fehr, Ernst and Fischbacher, Urs (2002) 'Why Do Social Preferences Matter – The Impact of Non-Selfish Motives on Competition, Cooperation and Incentives', *Economic Journal*, **112**, March, pp. C1-C33.

Fehr, Ernst and Gächter, Simon (2000a) 'Fairness and Retaliation: The Economics of Reciprocity', *Journal of Economic Perspectives*, **14**(3), Summer, pp. 159-81.

Fehr, Ernst and Gächter, Simon (2000b) 'Cooperation and Punishment in Public Goods Experiments', *American Economic Review*, **90**(4), December, pp. 980-995.

Fehr, Ernst and Gächter, Simon (2002) 'Altruistic Punishment in Humans', *Nature*, **415**, 10 January, 137-40.

Fehr, Ernst and Gintis, Herbert (2007) 'Human Motivation and Social Cooperation: Experimental and Analytical Foundations', *Annual Review of Sociology*, 33, pp. 43-64.

Field, Alexander J. (2001) *Altruistically Inclined? The Behavioral Sciences, Evolutionary Theory, and the Origins of Reciprocity* (Ann Arbor: University of Michigan Press).

Field, Alexander J. (2007) 'Beyond Foraging: Evolutionary Theory, Institutional Variation, and Economic Performance', *Journal of Institutional Economics*, **3**(3), December, pp. 265-291.

Fiske, Alan Page (1991) *Structures of Social Life: The Four Elementary Forms of Human Relations: Communal Sharing, Authority Ranking, Equality Matching, Market Pricing* (New York: Free Press).

Frey, Bruno S. (1997) *Not Just for Money: An Economic Theory of Personal Motivation* (Aldershot: Edward Elgar).

Ghiselin, Michael T. (1974) *The Economy of Nature and the Evolution of Sex* (Berkeley: University of California Press).

Gintis, Herbert (2000) 'Strong Reciprocity and Human Sociality', *Journal of Theoretical Biology*, 206, pp. 169-79.

Gintis, Herbert, Bowles, Samuel, Boyd, Robert and Fehr, Ernst (eds) (2005) *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life* (Cambridge, MA: MIT Press).

Greene, J. D. and Haidt, Jonathan (2002) 'How (and Where) Does Moral Judgement Work?' *Trends in Cognitive Sciences*, **6**(12), pp. 517-23.

Guzmán, Ricardo Andrés, Rodriguez-Sicken, Carlos and Rowthorn, Robert (2007) 'When in Rome, do as the Romans do: The Coevolution of Altruistic Punishment, Conformist Learning, and Cooperation', *Evolution and Human Behavior*, **28**, pp. 112-17.

Güth, Werner (1995) 'On Ultimatum Bargaining Experiments – A Personal Review', *Journal of Economic Behavior*, 27, pp. 329-344.

Haidt, Jonathan (2001) 'The Emotional Dog and the Rational Tail: A Social Interactionist Approach to Moral Judgment', *Psychological Review*, **108**, pp. 814-34.

Haidt, Jonathan and Joseph, Craig (2004) 'Intuitive Ethics: How Innately Prepared Intuitions Generate Culturally Variable Virtues', *Daedalus*, **133**(4), Fall, pp. 55-66.

Haidt, Jonathan and Joseph, Craig (2008) 'The moral mind: How five sets of innate intuitions guide the development of many culture-specific virtues, and perhaps even modules', in Peter Carruthers, S. Laurence, and S. Stich (eds.) *The Innate Mind, Vol. 3,* pp. 367-444.

Hamilton, William D. (1964) 'The Genetical Evolution of Social Behavior', *Journal of Theoretical Biology*, 7(1), July, pp. 1-32.

Hamlin, J. Kiley, Wynn, Karen and Bloom, Paul (2007) 'Social Evaluation by Preverbal Infants', *Nature*, vol. 450, pp. 557-560.

Hammerstein, Peter (ed.) (2003) *Genetic and Cultural Evolution of Cooperation* (Cambridge, MA: MIT Press).

Hare, Richard M. (1952) *The Language of Morals* (Oxford: Oxford University Press).

Hauser, Marc D. (2006) *Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong* (New York: Harper Collins).

Henrich, Joseph (2004) 'Cultural Group Selection, Coevolutionary Processes and Large-Scale Cooperation', *Journal of Economic Behavior and Organization*, **53**(1), February, pp. 3-35.

Henrich, Joseph and Boyd, Robert (1998) 'The Evolution of Conformist Transmission and the Emergence of Between Group Differences', *Evolution and Human Behavior*, **19**, pp. 215-42.

Henrich, Joseph, Boyd, Robert, Bowles, Samuel, Camerer, Colin, Fehr, Ernst, Gintis, Herbert and McElreath, Richard (2001) 'In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies', *American Economic Review (Papers and Proceedings),* **91**(2), May, pp. 73-84.

Henrich, Joseph, Boyd, Robert, Bowles, Samuel, Camerer, Colin, Fehr, Ernst, and Gintis, Herbert (2004) *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies* (Oxford and New York: Oxford University Press).

Henrich, Joseph, Jean Ensminger, Richard McElreath, Abigail Barr, Clark Barrett, Alexander Bolyanatz, Juan Camilo Cardenas, Michael Gurven, Edwins Gwako, Natalie Henrich, Carolyn Lesorogol, Frank Marlowe, David Tracer, and John Ziker, (2010) 'Markets, Religion, Community Size, and the Evolution of Fairness and Punishment', *Science*, 327 (5972), pp. 1480-84.

Henrich, Joseph and Gil-White, Francisco J. (2001) 'The Evolution of Prestige: Freely Conferred Deference as a Mechanism for Enhancing the Benefits of Cultural Transmission', *Evolution and Human Behavior*, **22**(3), pp. 165-96.

Henrich, Joseph, McElreath, Richard, Barr, Abigail, Ensminger, Jean, Barrett, Clark, Bolyanatz, Alaexander, Cardenas, Juan Camilo, Gurven, Michael, Gwako, Edwins, Henrich, Natalie, Lesorogol, Carolyn, Marlowe, Frank, Tracer, David, and Ziker, John (2006), 'Costly Punishment Across Human Societies'. *Science*, **312**, no. 5781, pp. 1767-1770.

Hill, Kim R., Robert S. Walker, Miran Božičević, James Eder, Thomas Headland, Barry Hewlett, A. Magdalena Hurtado, Frank Marlowe, Polly Wiessner, and Brian Wood (2011) 'Co-Residence Patterns in Hunter-Gatherer Societies Show Unique Human Social Structure', *Science*, 331, no. 6022, 11 March, pp. 1286-1289.

Hobson, John A. (1929) *Wealth and Life: A Study in Values* (London: Macmillan).

Hodgson, Geoffrey M. (1993) *Economics and Evolution: Bringing Life Back Into Economics* (Cambridge, UK and Ann Arbor, MI: Polity Press and University of Michigan Press).

Hodgson, Geoffrey M. (2007) 'Meanings of Methodological Individualism', *Journal of Economic Methodology* **14**(2), June, pp. 211-26.

Hodgson, Geoffrey M. (2009) 'On the Institutional Foundations of Law: The Insufficiency of Custom and Private Ordering', *Journal of Economic Issues*, **43**(1), March, pp. 143-66.

Hodgson, Geoffrey M. (2013) *From Pleasure Machines to Moral Communities: An Evolutionary Economics without Homo Economicus* (Chicago: University of Chicago Press).

Hodgson, Geoffrey M. and Knudsen, Thorbjørn (2010) *Darwin's Conjecture: The Search for General Principles of Social and Economic Evolution* (Chicago: University of Chicago Press).

Huxley, Thomas Henry and Huxley, Julian (1947) *Evolution and Ethics, 1893-1943* (London: Pilot Press).

Jevons, William Stanley (1871) *The Theory of Political Economy* (London: Macmillan).

Joyce, Richard (2006) *The Evolution of Morality* (Cambridge, MA: MIT Press).

Kaplow, Louis and Shavell, Steven (2007) 'Moral Rules, the Moral Sentiments, and Behavior: Toward a Theory of an Optimal Moral System', *Journal of Political Economy*, 115(3), June, pp. 494-514.

Kropotkin, Petr A. (1902) *Mutual Aid: A Factor of Evolution*, 1st edn. (London: Heinemann).

Lopes, Helena, Ana C. Santos and Nuno Teles (2009) 'The motives for cooperation in work organisations', *Journal of Institutional Economics*, **5**(3), December, pp. 259-88.

Mackie, John Leslie (1977) *Ethics: Inventing Right and Wrong* (Harmondsworth: Penguin).

Margolis, Howard (1982) *Selfishness, Altruism, and Rationality: A Theory of Social Choice* (Chicago: University of Chicago Press).

McCloskey, Deirdre N. (2006) *The Bourgeois Virtues: Ethics for an Age of Commerce* (Chicago: Chicago University Press).

McCrink, Koleen, Bloom, Paul and Santos, Laurie R. (2010) 'Children's and Adult's Judgments of Equitable Resource Distributions', *Developmental Science*, **13**(1), pp. 37-45.

Menger, Carl (1871) *Grundsätze der Volkwirtschaftslehre* (Tübingen: J. C. B. Mohr). Translated and published in English in 1981 as *Principles of Economics* (New York: New York University Press).

Midgley, Mary (2003) *Myths We Live By* (London and New York: Routledge).

Minkler, Lanse P. (2008) *Integrity and Agreement: Economics When Principles Also Matter* (Ann Arbor, MI: University of Michigan Press).

Nichols, Shaun (2004) *Sentimental Rules: On the Natural Foundations of Moral Judgment* (Oxford and New York: Oxford University Press).

Oppenheimer, Stephen (2004) *The Real Eve* (New York: Carroll and Graf). Published in the UK as *Out of Eden* (London: Constable and Robinson).

Palmer, Craig T. and Steadman, Lyle B. (1997) 'Human Kinship as a Descendant-Leaving Strategy: A Solution to an Evolutionary Puzzle' *Journal of Social and Evolutionary Systems*, **20**(1), pp. 39-51

Pinker, Steven (1994) *The Language Instinct: The New Science of Language and Mind* (London and New York: Allen Lane and Morrow).

Price, Michael E., Cosmides, Leda and Tooby, John (2002) 'Punitive Sentiment as an Anti-Free-Rider Device', *Evolution and Human Behavior*, **23**, pp. 203-31.

Rizzolatti, Giacomo, Corrado Sinigaglia and Frances Anderson (2007) *Mirrors in the Brain: How Our Minds Share Actions, Emotions, and Experience* (Oxford and New York: Oxford University Press).

Robinson, Paul H., Kurzban, Robert and Jones, Owen D. (2007) 'The Origins of Shared Intuitions of Justice', *Vanderbilt Law Review*, **60**(6), p. 1633-88.

Roberts, Simon (1979) *Order and Discipline: An Introduction to Legal Anthropology* (New York: St Martin's Press).

Runciman, Walter Garry (2002) 'Heritable Variation and Competitive Selection as the Mechanism of Sociocultural Evolution' in Wheeler, Michael, Ziman, John and Boden, Margaret A. (eds) (2002) *The Evolution of Cultural Entities* (Oxford: Oxford University Press), pp. 9-25.

Sampson, Geoffrey (2005) *The Language Instinct Debate* (London and New York: Continuum).

Schumpeter, Joseph A. (1942) *Capitalism, Socialism and Democracy*, 1st edn. (London: George Allen and Unwin).

Schwartz, Shalom H. (1994) 'Are There Universal Aspects in the Structure and Contents of Human Values?' *Journal of Social Issues*, **50**(4), pp. 19-45.

Sen, Amartya K. (1987) *On Ethics and Economics* (Oxford and New York: Basil Blackwell).

Smart, J. J. C. and Williams, Bernard (1973) *Utilitarianism: For and Against* (Cambridge: Cambridge University Press).

Sober, Elliott and Wilson, David Sloan (1998) *Unto Others: The Evolution and Psychology of Unselfish Behavior* (Cambridge, MA: Harvard University Press).

Tancredi, Laurence (2005) *Hardwired Behavior: What Neuroscience Reveals about Morality* (Cambridge and New York: Cambridge University Press).

Thiroux, Jacques P. and Krasemann, Keith W. (2009) *Ethics: Theory and Practice,* 10th edn. (London: Pearson).

Tisak, Marie S. and Turiel, Elliot (1984) 'Children's Conceptions of Moral and Prudential Rules', *Child Development,* **55**, pp. 1030-39.

Trivers, Robert L. (1971) 'The Evolution of Reciprocal Altruism', *Quarterly Review of Biology,* **46**(1), pp. 35-57.

Vanberg, Viktor J. (2008) 'On the Economics of Moral Preferences', *American Journal of Economics and Sociology*, **67**(4), October, pp. 605-28.

Wade, Michael J. (1978) 'A Critical Review of the Models of Group Selection', *Quarterly Review of Biology*, **53**, pp. 101-14.

Walras, Léon (1874) *Éléments d'économie politique pure, ou théorie de la richesse sociale* (Lausanne : Rouge).

Wiessner, Polly (2005) 'Norm Enforcement among the Ju/'hoansi Bushmen: A Case of Strong Reciprocity?' *Human Nature,* 16(2), pp. 115-45.

Westermarck, Edward (1891) *The History of Human Marriage* (London and New York: Macmillan).

Wicksteed, Philip H. (1910) *The Commonsense of Political Economy, including a Study of the Human Basis of Economic Law* (London: Macmillan).

Williams, George C. (1966) *Adaptation and Natural Selection* (Princeton, NJ; Princeton University Press).

Wilson, David Sloan (1980) *The Natural Selection of Populations and Communities* (Menlo Park, CA: Benjamin/Cummings).

Wilson, David Sloan (1983) 'The Group Selection Controversy: History and Current Status', *Annual Review of Ecology and Systematics*, **14**, pp. 159-88.

Wilson, David Sloan (2002) *Darwin's Cathedral: Evolution, Religion, and the Nature of Society* (Chicago: University of Chicago Press).

Wilson, David Sloan and Sober, Elliott (1994) 'Reintroducing Group Selection to the Human Behavioral Sciences', *Behavioral and Brain Sciences*, **17**(4), pp. 585-608.

Wilson, David Sloan and Wilson, Edward O. (2007) 'Rethinking the Theoretical Foundations of Sociobiology', *Quarterly Review of Biology,* 82(4), December, pp. 327-48.

Zak, Paul J. (2004) 'Neuroeconomics', *Philosophical Transactions of the Royal Society of London,* **B**, **359**, pp. 1737-48.

Zak, Paul J. (ed.) (2008) *Moral Markets: The Critical Role of Values in the Economy* (Princeton: Princeton University Press).