BOOK REVIEW FEATURE: TWO BOOK REVIEWS OF *A COOPERATIVE SPECIES: HUMAN RECIPROCITY AND ITS EVOLUTION*

By Bowles (Samuel) and Gintis (Herbert) (Princeton, NJ: Princeton University Press, 2011, pp. 288, $39.95/£27.95, cloth, ISBN: 9780691151250.

Review 1

One day, wandering south along the continental divide of North America, I came alongside a traveller from Chicago. He told me one should never argue about the Rocky Mountains. 'They simply exist', he said, 'constant and immutable'. I thanked him for his insight and continued on my way, happy to have one less thing to argue about. Somewhere near Santa Fe, I came alongside another traveller with an interest in geophysics. To my surprise, he spoke endlessly about the Rocky Mountains, why they are as they are and how they constantly change. We talked for many hours, even arguing one might say, about this theory and that theory, about the evidence for and the evidence against. For him, the mountains were neither fixed nor axiomatic. They could have been another way, but they were not, at least not for the moment. By his way of thinking, the Rocky Mountains were endogenous, and their origin and evolution was the puzzle that absorbed him.

Never ones to shy away from a healthy scientific controversy, Sam Bowles and Herb Gintis are arguing about the Rocky Mountains. Specifically, if we drop the Rocky Mountain analogy (Stigler and Becker, 1977), they are arguing about the evolutionary origin of human altruism and the social preferences that sustain this distinctive form of pro-social behaviour. In a narrow sense, consequently, their most recent book is not an economics book. Although Bowles and Gintis clearly have an economist's interest in social policy, and although they often position their arguments relative to big questions in the history of economic thought, their scientific task takes them well beyond the traditional methods and conventional topics of economics.

In a broad sense, however, *A Cooperative Species* is clearly a book about economics because Bowles and Gintis examine fundamental questions about human behaviour and its relation to the distribution of resources in finite environments. To do so, they work with an expanded repertoire of questions and methods, and the need for this expanded repertoire can be summarised as follows. The last two or three decades have brought a considerable body of research showing that people often exhibit strong forms of altruism like cooperating in anonymous one-shot interactions with genetically unrelated individuals (Camerer, 2003; Fehr and Fischbacher, 2003; Henrich *et al.*, 2006, 2010). This finding is controversial but the evidence continues to accumulate. For present purposes, the implication is straightforward. If one takes an abiological social science perspective, like that of economics, the observation that people behave

altruistically can potentially be readily accommodated. So long as the behaviour in question is consistent, say because it satisfies the axioms of revealed preference (Andreoni and Miller, 2002), one simply needs to find the best or perhaps the most useful utility representation and one is off and running. Indeed, a number of researchers have taken up exactly this task (Rabin, 1993; Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Charness and Rabin, 2002; Falk and Fischbacher, 2006).

From a biological perspective, however, the matter cannot be reduced to one of exogenous preferences. One must instead explain how a population of material organisms who live in a finite environment and reproduce biologically, which is a material phenomenon, can evolve in a way that leads individuals to pay a material cost to produce material benefits for others. Social preferences may be one type of important and fascinating explanation for altruistic behaviour, but delineating the subtleties of social preferences will not address the question of where these preferences came from in the first place (Boyd and Richerson, 2005). As Bowles and Gintis point out, economists have often justified the self-interest assumption conventional in economics with off-the-cuff arguments about how natural selection must produce selfish organisms. The overarching point of Bowles and Gintis is to challenge this loose argument. Their book is an extended analysis showing that, not only is the self-regarding assumption inconsistent with empirical evidence, it is also *not* a necessary consequence of natural selection. In short, 'the idea that selfish genes must produce selfish individuals is false' (p. 45).

As we will see below, their counterargument centres on the fact that humans live in groups of a very special sort. Namely, human social groups are the loci of culturally evolved social norms and institutions that moderate selfish behaviour within groups. Across group boundaries, however, people are often selfish, parochial and mutually hostile. Both of these group-level attributes can increase selection for altruism that is, in one way or another, conditional on group affiliation. This happens because of what is called 'multilevel selection'. Before we get to that, however, I would like to turn to the Pleistocene because it is a difficult and recurring problem for those who do research on the evolution of human social behaviour. The rigour and insight Bowles and Gintis demonstrate when managing the Pleistocene problem also happens to be a large part of why I admire their book so much.

1. Methodologies for the Pleistocene Problem

The late Pleistocene is tricky for two basic reasons. First, it was sufficiently long ago that it was before people started writing things down, and so we have a limited understanding of what it was like in terms of human social behaviour and the preferences, beliefs and constraints underlying this behaviour. Second, many distinctive features of human cognition, including those related to human altruism, probably evolved in the late Pleistocene when all humans lived as hunter-gatherers in small-scale societies. An evolutionary process of critical scientific interest is thus not available for direct observation; only the outcome is available. In general, this difficulty is intrinsic to the evolutionary study of human social behaviour. It makes scientific investigation hard, but not impossible. Many scientific disciplines take long-term historical processes as objects of study. Our metaphorical geophysicist from above comes to mind. In addition, many

scientists study contemporary phenomena at a scale (e.g. molecular biology) that makes observation, at the very least, complicated. Researchers nonetheless make steady progress using a combination of theoretical and empirical methods. In such cases, an important question centres on exactly how to proceed methodologically when addressing phenomena that are difficult or impossible to observe.

The evolutionary study of human behaviour has not always made stunning progress on this front. Human behavioural ecology has made admirable contributions on important questions like how people forage, whom they mate with and how much they invest in their children (Winterhalder and Smith, 1992; Cronk *et al.*, 2000). The study of gene–culture coevolution (Cavalli-Sforza and Feldman, 1981; Boyd and Richerson, 1985; Durham, 1991; Richerson and Boyd, 2005; Rendall *et al.*, 2010), which plays a critical role in Bowles and Gintis' account of human altruism, has long been a theoretical undertaking and an empirical programme has recently begun to emerge (Smith *et al.*, 2008). All in all, however, evolutionary ecologists have often struggled with human behaviour, especially human social behaviour, and the field is vulnerable to the criticism that it involves a considerable amount of story telling about, for example, life in the Pleistocene.

Like many researchers, Bowles and Gintis confront the Pleistocene problem directly by positing a particular view of life in the Pleistocene and analysing the implications. In this sense, their methodology is consistent with that of others. What distinguishes their work is their ability to draw from a wide variety of sources and integrate the resulting information in a compelling and quantitatively precise way. To a large extent, Bowles and Gintis work with models. They construct and analyse their models, however, by integrating theory and data from many fields, including behavioural economics, population genetics, archaeology, ethnography and human evolutionary ecology. They excel, quite simply, at incorporating information from many scientific corners. This is not interdisciplinary research for its own sake; it is interdisciplinary research for the sake of tackling a highly controversial scientific problem that necessarily spans numerous disciplines (Pennisi, 2005).

Other researchers specialising in the evolution of human altruism also make claims about the Pleistocene. Evolutionary psychologists constitute an especially active group in this category. Evolutionary psychologists generally begin their analyses with the notion that modern human cognition evolved in the late Pleistocene, and it is thus adapted to life as part of a small band of hunter-gatherers (Smith, 2000). At this general level, the point is relatively innocuous. The way most people live now is evolutionarily recent. Corn and other domesticated grasses, high population densities, economic inequality, antibiotics, well-developed markets, pigs and other animal domesticates are all new developments. However quickly natural selection might be able to reorganise human decision making to deal with these modern innovations, a lag certainly exists, and so at least some of how we make decisions now is probably the residue of the late Pleistocene. This simple idea has come to play a central role in the debate over the evolutionary origins of human altruism (Hagen and Hammerstein, 2006).

For example, evolutionary psychologists have consistently argued that strong altruistic motives have little importance (Haley and Fessler, 2005). They argue instead that social interactions in the late Pleistocene were always repeated and always occurred with either close genetic relatives or with members of small, stable social groups. As a result,

any seemingly altruistic behaviour that might have occurred was simply a product of what Bowles and Gintis call 'enlightened self-interest'. People helped their kin, and they helped others in ways that maintained their reputations as good members of the social group. This latter reputational effect, in turn, ensured that the helpers would themselves receive help from others in future interactions. The mechanisms here are straightforward processes associated with kin selection (Hamilton, 1964) and reciprocity (Trivers, 1971; Nowak and Sigmund, 1998), both of which are evolutionary mechanisms producing specific types of pro-social behaviour that are relatively easy to explain. The critical claim of evolutionary psychologists is that our cognition is still adapted to kin groups and other persistent social contacts like those characteristic of the late Pleistocene. Deep within the recesses of our hunter-gatherer minds, we do not really understand anonymous one-shot interactions with unrelated individuals, interactions that typify experiments on social preferences and provide much of the evidence for strong forms of human altruism that are hard to explain.

Even though this argument has been very influential, the conclusions do not follow from the evidence. The evidence tends to centre on the experimental finding that payoff-irrelevant social cues, like a pair of stylised eyes on an experimental subject's computer screen, increase altruistic giving in dictator games (Haley and Fessler, 2005; Rigdon *et al.*, 2009). In other words, an experimentalist can modulate altruism with cues that potentially make subjects feel like they are being watched but do not actually affect payoffs in any way. The link with the Pleistocene is the following. Eyespot results, the argument goes, derive from the fact that these irrelevant cues activate an ancestral hunter-gatherer psychology sensitively attuned to the possibility that one is being watched and that one's reputation is at stake. The altruism observed in economic experiments then is not really altruism; it is a cognitive distortion. Every experiment has payoff-irrelevant social cues (e.g. multiple subjects playing at once on a local computer network) and, as a result, experiments that have an explicit structure that is anonymous and one-shot have an implicit structure, in the hunter-gatherer minds of university undergraduates, that is associated with repeated interactions and reciprocity. This may sound far-fetched but eyespot results have proven replicable with simple allocation games like the dictator game. Nonetheless, as Bowles and Gintis argue, the fact that a researcher can modulate pro-social behaviour by providing cues that a person is being observed does not logically imply that all pro-social behaviour can or should be explained in terms of some unconscious activation of a hunter-gatherer's concern with his reputation. Indeed, the experimental evidence is unambiguous. Subjects do distinguish between anonymous one-shot interactions and repeated interactions with *explicit* reputational mechanisms. Moreover, when we move beyond dictator games, subjects do not reliably respond to eyespots (Fehr and Schneider, 2010).

The interpretation of eyespots is an example of the big-mistake hypothesis. Apart from the multilevel selection account of Bowles and Gintis, the big-mistake hypothesis is the primary alternative explanation for strong forms of altruism. Under the big-mistake hypothesis, human altruism is a big mistake in evolutionary terms because a decision-making process evolved to deal with an ancestral environment is misapplied in contemporary settings. The mistake occurs because the evolution of human social cognition has not caught up with the novel features of contemporary social life. This is

a fine idea, but we have no particular reason to believe that the evolutionary lag is such that cognition specifically reflects the repeated interactions experienced by those who lived in the late Pleistocene. Indeed, many evolutionary ecologists who study human behaviour disagree strongly with this notion and think the case for lags of this sort is overstated (Smith *et al.*, 2011). Moreover, even if the lag places modern cognition squarely in the late Pleistocene, contemporary decision making could still reflect some feature of ancestral social life that did not involve repeated interactions. This theoretical possibility leaves us with our original challenge of characterising human social life in the Pleistocene. In any case, the justification for human altruism as an expression of repeated interactions in the past is often a justification based on exclusion. Without a big mistake of this sort, we are left with some form of multilevel selection in the spirit of Bowles and Gintis. Multilevel selection has a long and controversial history in evolutionary ecology. In particular, selection at the level of the social group is sufficiently problematic to be considered an extraordinary explanation for human social behaviour (Haley and Fessler, 2005). One can thus argue for big mistakes related to repeated interactions in the past, with all the necessary assumptions about evolutionary lags and social life in the Pleistocene, by simply pointing out that these big mistakes do not require multilevel selection.

The account of Bowles and Gintis relies heavily on multilevel selection, and they could never justify their arguments by saying these arguments do not require kin selection or repeated interactions. Even though humans behave altruistically in anonymous one-shot interactions with non-kin, kin selection and repeated interactions are still offered as explanations by many researchers. The standard of evidence required of Bowles, Gintis and their colleagues is thus higher than that required of those who explicitly argue for big mistakes. Much of what I value most in Bowles and Gintis is that they do not shy away from this asymmetric standard of evidence. They simply go. Their view of the Pleistocene is different from proponents of a big mistake. They argue that ancestral humans most likely had many evanescent human contacts, and ancestral humans most likely lived in social groups too large, too fluid and too genetically unrelated for reciprocity and kinship alone to sustain pro-social behaviour. In the process of developing this view, they use genetic data, paleoclimatology and the ethnographic record. They insist on the discipline of both models and data from all relevant sources and, by doing so, they make the critical issues clearer and more precise than vague claims about evolutionary lags, the reorganisation of human cognition, misfiring preferences and distorted beliefs. Whatever your final take on the specific arguments of Bowles and Gintis, whatever you think of their view of the archaeological record from the Pleistocene or however you judge their interpretation of experimental evidence, they deserve tremendous credit for pushing the evolutionary approach to human social behaviour forward by insisting on an integrated approach to theory and empiricism. The result is a degree of rigour that is not as common as it should be but it is arguably essential when trying to understand an evolutionary process that has already occurred.

To do all this, Bowles and Gintis use two basic approaches. First, they marshal evidence from archaeology, population genetics, paleoclimatology and ethnographies to develop as precise a view of ancestral human social life as they can. Bowles and Gintis are especially interested in who interacted with whom, what cooperation meant in practice, how often people died from inter-group warfare and how volatile the

environment was. The answers to these questions help them to identify regions of interest in the parameter spaces of their models. Knowing where to go in parameter space is always nice, of course, but when the evolution of human altruism is concerned, it takes on a special importance. When modelling complex evolutionary processes in a population subdivided into groups, analytical methods are often limited. Simulations are frequently irreplaceable (Bowles, 2004) and so placing bounds on parameter values can be crucial. Bowles and Gintis use both analytical and simulation models to examine the evolution of human altruism *in the relevant region of parameter space*. Concentrating on the relevant conditions is obvious, important and, unfortunately, rare. We know little, for example, about how big a big mistake has to be and whether a cognitive distortion of the required magnitude is in any way reasonable.

The second basic methodological approach of Bowles and Gintis depends on proximate mechanisms. In evolutionary biology, one often distinguishes between proximate and ultimate causation. Ultimate causes refer to evolutionary function. For example, I eat because in ancestral human populations eating resulted in higher fitness on average than not eating. Proximate causes refer to mechanisms that can be analysed and understood without recourse to evolutionary function. I eat, for example, because I am hungry. To put the matter in economic terms, preferences and beliefs can be thought of as proximate, while the evolutionary process that would lead a population to have an observed distribution of preferences and beliefs is ultimate. With respect to human altruism, proximate motivations can potentially be measured via careful behavioural analyses, well-designed questionnaires, brain imaging methods and per- haps other techniques that await development. Because the proximate motivations for contemporary behaviour represent the residue of our evolution, insofar as different evolutionary theories of human altruism imply different proximate mechanisms, the careful study of proximate mechanisms can illuminate the past. If two evolutionary theories of human altruism predict distributions of preferences and beliefs that are not identical, we have hope. The right kinds of measurements on contemporary behaviour and its underlying motives can potentially discriminate between the theories. Bowles and Gintis develop this method, especially when they consider the processes by which people are socialised into the cultural and normative systems of their societies. For example, they review evidence indicating that people adhere to social norms like those promoting fairness because the norms themselves become internalised. In other words, norm adherence leads to trouble-free social interactions, and so it has extrinsic value. When internalised, norms become a part of the preference relation itself, and so norm adherence also has intrinsic value. After reviewing the evidence that humans internalise fairness norms, they go on to show how the capacity to internalise norms can evolve in theory. They further show that internalisation, once in place, can sustain individually costly but group-beneficial norms like altruism. Importantly, these processes are endogenous, and as an illustrative example the evolution of internalised norms rep- resents novel and important methodological territory. My summary is the following. The multilevel selection arguments of Bowles, Gintis and their colleagues are contro- versial; they rely on forms of selection that are in many ways heterodox in evolutionary ecology. Precisely to confront the controversy, however, they have pushed the standard of evidence far forward. Even apart from its many other virtues, *A Cooperative Species* deserves a wide readership for this reason alone.

2. The Multilevel Selection of Group-living, Norm-abiding, Warrior Altruists

Biological systems are hierarchically organised. Genes are grouped into chromosomes. Chromosomes are grouped into cells. Cells are grouped into individual organisms. Individual organisms are grouped into social groups. Given this kind of hierarchical organisation, the total selective pressure in a population can be decomposed into selection at every hierarchical level (Price, 1970). This is what is called multilevel selection (Keller, 1999). Simply put, altruism towards genetically unrelated individuals is a difficult puzzle because the altruist pays an individual cost to produce a positive externality that obtains at the group level. This kind of social interaction is not generally favourable for the genetic evolution of altruism because selection on genes at the individual level will almost always dominate selection at the group level (Williams, 1966). Specifically, two quantities matter at each level. First, selection at the level in question depends on the variance in traits (e.g. altruism versus selfishness) at that level. Intuitively, this is true because selection, natural or otherwise requires variation. Second, selection also depends on the relationship between fitness at the level in question and the prevalence of a particular trait like altruism. Altruists have lower fitness than selfish individuals, and so selection is against altruism at the individual level. The relationship between altruism and individual fitness is negative. However, because of the efficiency gains associated with altruism, groups with many altruists have higher average fitness than groups with few altruists. Thus, altruistic groups can potentially do better than selfish groups when groups compete. The relationship between altruism and fitness is positive at the group level. Individual selection is thought to typically dominate selection among groups because small amounts of gene flow among groups tend to dramatically reduce the variance among groups (Roughgarden, 1979). Given limited variation at the group level, the efficiency gains from altruism would typically have to be extraordinary for the group-level benefits to offset the individual-level costs. This is why the evolution of altruism is one of the big unsolved problems in contemporary science (Pennisi, 2005).

A *Cooperative Species* is an extended account of why the reasoning above is not adequate for humans. In terms of proximate mechanisms, the thesis of Bowles and Gintis rests on the presence of social preferences, social emotions such as shame and guilt, and the psychological tendency to internalise social norms promoting pro-social behaviour. In terms of ultimate mechanisms, their argument rests on war and other forms of hostility among groups, volatile environments and the gene–culture coevolution of norms and institutions that reduce variation in fitness within groups and maintain variation between groups. In principle, all these ultimate mechanisms should increase the importance of selection at the level of the social group and, in this sense, they should all improve conditions for the evolution of proximate mechanisms that support individually costly but group-beneficial behaviours.

Gene–culture coevolution (Richerson and Boyd, 2005) plays an especially critical role here. Gene–culture coevolution refers to the fact that human evolution rests on two systems of inheritance, namely genes and culture. Specifically, humans transmit information that influences behaviour both genetically and via social learning and cultural transmission. Our genes influence our cognition and by extension the way cultural information is transmitted, while culturally transmitted information can lead

humans to dramatically restructure their physical and social environments in ways that change the selective pressures on genes (Olding-Smee *et al.*, 2003). The result is a coevolutionary system governed by linked dynamical processes in both genetic and cultural dimensions (Boyd and Richerson, 1985; Richerson and Boyd, 2005). Gene–culture coevolution has the interesting property that many of the hypothesised transmission mechanisms tend to reduce behavioural variation within groups and maintain the variation that exists between groups. This has long been a theoretical point but the pattern also fits nicely with an intuitive understanding of what we mean by a cultural group. Moreover, empirical results validating such patterns are accumulating (Richerson and Boyd, 2005; Bell *et al.*, 2009), including some beautiful results showing how cultural variation in preferences and norms can affect labour markets (Bruegger *et al.*, 2009). The upshot is persistent behavioural variation between groups and, all else equal, increased scope for selection at the group level.

More to the point, Bowles and Gintis concentrate heavily on the cultural evolution of norms and institutions that exaggerate the importance of selection among groups, the level at which selection favours altruism, and attenuate selection within groups, the level at which selection favours selfishness. In the former category, war and other forms of inter-group conflict can intensify the selective pressure favouring altruism at the group level. Specifically, altruism generates efficiency gains and so groups with many altruists will be especially productive. When groups come into conflict, the more productive groups with more altruists should do better on average. Bowles and Gintis estimate that conflicts of this sort should have taken two basic forms, one endogenous to their theory and the other exogenous. War is the endogenous source of conflict and one of the key models in their book shows how out-group hostility and in-group altruism coevolve. Both types of behaviour are mutually reinforcing because each improves conditions for the evolution of the other. As is typical of their method, Bowles and Gintis do not simply make this theoretical point. They go to considerable lengths to show that both the assumptions and conclusions of the model are consistent with what we know from the archaeological and ethnographic records about life in ancestral societies. Our ancestors, in short, were impressively warlike. Environmental instability is the exogenous source of conflict. The idea here is that, even without open war, groups can come into strong if indirect conflict as they vie for resources in volatile, limited environments. Again, Bowles and Gintis draw from recent research in paleoclimatology to justify this idea. In this case, the research clearly shows that the Pleistocene was a dramatic time to live. The environments in which modern humans evolved were characterised by stunning volatility. This presumably led to the on-going reorganisation of ecological communities and, by extension, chronic difficulties for ancestral hunter-gatherers as they attempted to acquire the biological resources essential to their survival.

Within groups, Bowles and Gintis examine the evolution of several institutions that reduce the force of selection against altruistic individuals. These institutions specifically include punishment, reproductive levelling and segmentation. Punishment has received a considerable amount of theoretical and empirical attention (Boyd and Richerson, 1992; Fehr and Gaechter, 2002; Gaechter *et al.*, 2008). Conditional on the presence of people willing to punish bad behaviour, punishment can induce self-regarding individuals to cooperate. This can limit the selective disadvantage

experienced by altruists within groups. If costly, however, punishment is itself often a form of altruism, and so the evolution of punishment is a second-order theoretical puzzle. Moreover, punishment destroys resources, and if used heavily, it can reduce the advantages of cooperation at the group level. Bowles, Gintis and their colleagues have been actively involved in solving these puzzles for many years (Boyd and Richerson, 1992; Boyd *et al.*, 2003, 2010). Their book presents the culmination of these efforts with a model showing how punishment can evolve endogenously as a socially coordinated activity that stabilises altruism within a social group. Again, as always, Bowles and Gintis make special efforts to show how the model is consistent with what we know about punishment in small-scale societies.

Like punishment, reproductive levelling and segmentation are also mechanisms that should restrict the disadvantages of altruism within a social group. Reproductive levelling refers to norms and institutions that promote the redistribution of resources within the group. Food sharing is an well-documented and important example of levelling in small-scale societies (Boehm, 1999; Kaplan and Gurven, 2005). Segmentation relates to the process by which people are matched to interact socially. In particular, if social interactions take the form of a social dilemma, segmentation typically refers to a non-random matching process that partially sequesters altruists from exploitation by non-altruists. For example, consider a group with a proportion, $p \in (0,1)$, of altruists. If segmentation holds, altruists interact with altruists with some probability greater than p, while self-regarding individuals interact with altruists with some probability less than p. These differential matching probabilities mean that altruists are shielded, statistically speaking, from exploitation. Segmentation can occur for many reasons ranging from simple biological kinship to social institutions that publicise an individual's reputation and so allow others to interact conditionally with the individual in question. Importantly, if segmentation holds, the risk of exploitation for altruists is lower than it is under random matching. For this reason, segmentation moderates the selection against altruism at the individual level. Indeed, Bowles and Gintis show that even small amounts of segmentation can dramatically increase the region of parameter space for which selection among groups promotes the evolution of altruism. As with punishment, Bowles and Gintis show how levelling and segmentation can evolve endogenously under conditions estimated to be characteristic of the Pleistocene.

Altogether, Bowles and Gintis present a number of models, each of which isolates certain key mechanisms related to the evolution of human altruism. The models detail a species that lives in groups of an unusual kind. These groups consist of individuals with a remarkable ability to develop and acquire skills, beliefs and modes of social organisation via social learning. The result is the cultural evolution of norms and institutions that become group-level properties. If some groups have norms and institutions promoting war between groups along with punishment, levelling and segmentation within groups, then the selective regime is especially suitable for the evolution of individually costly, group-beneficial behaviour. We label these types of behaviour altruism. Importantly, however, the question of why other species have not evolved in a similar way receives limited attention by Bowles and Gintis. Many species learn socially and construct physical and social environments, though not with the facility of humans. Many species also exhibit out-group hostility, though again without

the regrettable alacrity of humans. Bowles and Gintis suggest that the remarkable abilities of humans in multiple domains came together to create a unique synergy favouring the evolution of altruism. This may be true, but this difficult question remains open. On the one hand, humans may not be unique in terms of exhibiting strong forms of altruism (Burkart *et al.*, 2007). On the other hand, even if humans are unique in this respect, the general mechanisms examined by Bowles and Gintis do not seem especially distinctive when considered in isolation. As a consequence, some kind of synergy between mechanisms is an important possibility requiring further research. In the end, all theories of human altruism struggle to some extent with identifying how distinctive humans are, in what ways exactly, and why. Even big mistakes, which provide an uncomplicated explanation of human altruism, provide no insight about why humans make such mistakes, while chimpanzees do not (Silk *et al.*, 2005; Jensen *et al.*, 2007), and common marmosets do (Burkart *et al.*, 2007).

We now have several theories for the evolution of pro-social behaviour, including strong forms of altruism. Bowles and Gintis do not try to identify which theory works. As they argue, all candidate theories work under the right conditions. Bowles and Gintis instead do the following. First, they develop a theory that implies decision-making processes consistent with their view of how modern humans make decisions in social settings. This task is difficult because it asks us to identify, in addition to the decisions themselves, the processes behind them. As a result, a reader firmly committed to the revealed preference approach may not disagree with Bowles and Gintis but he or she will find some parts frustrating. Second, if all models work, Bowles and Gintis want to identify which one can produce the evolution of altruism, as it exists in modern humans, under the relevant ancestral conditions. Specifying the relevant ancestral conditions is an intricate undertaking plagued by uncertainty. Our understanding of the Pleistocene is underdetermined. The greatest scope for disagreement with Bowles and Gintis lies here, and many researchers will no doubt disagree. Nonetheless, Bowles and Gintis have written an outstanding book that presents an important contribution and quite simply raises the scientific standard associated with the difficult and contentious problem of how human altruism evolved.

CHARLES EFFERSON

Department of Economics, University of Zürich

References

Andreoni, J. and Miller, J. (2002). 'Giving according to GARP: an experimental test of the consistency of preferences for altruism', *Econometrica*, vol. 70(2), pp. 737–53.

Bell, A.V., Richerson, P.J. and McElreath, R. (2009). 'Culture rather than genes provides greater scope for the evolution of large-scale human prosociality', *Proceedings of the National Academy of Sciences*, vol. 106(42), pp. 17671–4.

Boehm, C. (1999). *Hierarchy in the Forest: The Evolution of Egalitarian Behavior*, Cambridge, MA: Harvard University Press.

Bolton, G.E. and Ockenfels A. (2000). 'ERC: a theory of equity reciprocity and competition', *American Economic Review*, vol. 90(1), pp. 166–93.

Bowles, S. (2004). *Microeconomics: Behavior Institutions and Evolution*, New York: Russell Sage.

Boyd, R. and Richerson, P.J. (1985). *Culture and the Evolutionary Process*, Chicago, IL: University of Chicago Press.

Boyd, R. and Richerson P.J. (1992). 'Punishment allows the evolution of cooperation (or anything else) in sizable groups', *Ethology and Sociobiology*, vol. 13, pp. 171–95.

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

Boyd, R., Gintis, H. and Bowles S. (2010). 'Coordinated punishment of defectors sustains cooperation and can proliferate when rare', *Science*, vol. 328, pp. 617–20.

Boyd, R., Gintis, H., Bowles, S. and Richerson, P.J. (2003). 'The evolution of altruistic punishment', *Proceedings of the National Academy of Sciences*, vol. 100(6), pp. 3531–5.

Bruegger, B., Lalive, R. and Zweimueller, J. (2009). Does culture affect unemployment? evidence from the Roestigraben, I.Z.A. Discussion Paper No. 4283, University of Lausanne.

Burkart, J., Fehr, E., Efferson, C., and van Schaik, C.P. (2007). 'Other-regarding preferences in a non-human primate: common marmosets provision food altruistically', *Proceedings of the National Academy of Sciences*, vol. 104(50), pp. 19762–6.

Camerer, C.F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton: Princeton University Press.

Cavalli-Sforza, L.L. and Feldman, M.W. (1981). *Cultural Transmission and Evolution: A Quantitative Approach*, Princeton: Princeton University Press.

Charness, G. and Rabin, M. (2002). 'Understanding social preferences with simple tests', *Quarterly Journal of Economics*, vol. 117(3) pp. 817–69.

Cronk, L., Chagnon, N. and Irons, W., (eds.) (2000). *Human Behavior: An Anthropological Perspective*, New York: Aldine de Gruyter.

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

Fehr, E. and Fischbacher, U. (2003). 'The nature of human altruism', *Nature*, vol. 425(23 October), pp. 785–91.

Falk, A. and Fischbacher, U. (2006). 'A theory of reciprocity', *Games and Economic Behavior*, vol. 54(2), pp. 293–315.

Fehr, E. and Gaechter, S. (2002). 'Altruistic punishment in humans', *Nature*, vol. 415(10 January), pp. 137–40.

Fehr, E. and Schmidt, K.M. (1999). 'A theory of fairness competition and cooperation', *Quarterly Journal of Economics*, vol. 114(3), pp. 817–68.

Fehr, E. and Schneider, F. (2010). 'Eyes are on us, but nobody cares: are eye cues relevant for strong reciprocity?', *Proceedings of the Royal Society B: Biological Sciences*, vol. 277(1686), pp. 1315–23.

Gaechter, S., Renner, E. and Sefton, M. (2008). 'The long-run benefits of punishment', *Science*, vol. 322(5 December), p. 1510.

Hagen, E.H. and Hammerstein, P. (2006). 'Game theory and human evolution: a critique of some recent interpretations of experimental games', *Theorectical Population Biology*, vol. 69(3), pp. 339–48.

Haley, K.J. and Fessler, D.M.T. (2005). 'Nobody's watching? Subtle cues affect generosity in an anonymous economic game', *Evolution and Human Behavior*, vol. 26(3), pp. 245–56.

Hamilton, W.D. (1964). 'The genetical evolution of social behavior', *Journal of Theoretical Biology*, vol. 7(1), pp. 1–52.

Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., Cardenas, J.C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D. and Ziker, J. (2010). 'Markets, religion, community size, and the evolution of fairness and punishment', *Science*, vol. 327(5972), pp. 1480–4.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J.C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D. and Ziker J. (2006). 'Costly punishment across human societies', *Science*, vol. 312(23 June), pp. 1767–70.

Jensen, K., Call, J. and Tomasello, M. (2007). 'Chimpanzees are rational maximizers in an ultimatum game', *Science*, vol. 318(5 October), pp. 107–9.

Kaplan, H. and Gurven, M. (2005). 'The natural history of human food sharing and cooperation: a review and a new multi-individual approach to the negotiation norm', in (H. Gintis, S. Bowles, R. Boyd and E. Fehr, eds.), *Moral Sentiments and Material Interests: The Foundation of Cooperation in Economic Life*, pp. 75–113. Cambridge, MA: The MIT Press.

Keller, L., (ed.) (1999). *Levels of Selection in Evolution*, Princeton: Princeton University Press.

Nowak, M.A. and Sigmund, K. (1998). 'The dynamics of indirect reciprocity', *Journal of Theoretical Biology*, vol. 194(4), pp. 561–74.

Olding-Smee, F. J., Laland, K.N. and Feldman, M.W. (2003). *Niche Construction: The Neglected Process in Evolution*, Princeton: Princeton University Press.

Pennisi, E. (2005). 'How did cooperative behavior evolve?', *Science*, vol. 309, p. 93.

Price, G.R. (1970). 'Selection and covariance', *Nature*, vol. 227, pp. 520–1.

Rabin, M. (1993). 'Incorporating fairness into game theory and economics', *American Economic Review*, vol. 83(5), pp. 1281–1302.

Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M.W., Fogarty, L., Ghirlanda, S., Lillicrap, T. and Laland, K.N.(2010). 'Why copy others? Insights from the social learning strategies tournament', *Science*, vol. 328(5975) pp. 208–13.

Richerson, P.J. and Boyd, R. (2005). *Not By Genes Alone: How Culture Transformed the Evolutionary Process*, Chicago, IL: University of Chicago Press.

Rigdon, M., Ishii, K., Watabe, M. and Kitayama, S. (2009). 'Minimal social cues in the dictator game', *Journal of Economic Psychology*, vol. 30(3), pp. 358–67.

Roughgarden, J. (1979). *Theory of Population Genetics and Evolutionary Ecology: An Introduction*, Upper Saddle River, NJ: Prentice Hall.

Silk, J.B., Brosnan, S.F., Vonk, J., Henrich, J., Povinelli, D.J., Richardson, A.S., Lambeth, S.P., Mascaro, J. and Schapiro, S.J. (2005). 'Chimpanzees are indifferent to the welfare of unrelated group members', *Nature*, vol. 437(27 October), pp. 1357–9.

Smith, E.A. (2000). 'Three styles in the evolutionary analysis of human behavior', in (L. Cronk, N. Chagnon and W. Irons eds.), *Adaptation and Human Behavior: An Anthropological Perspective*, pp. 27–46. New York: Aldine de Gruyter.

Smith, E.A., Borgerhoff Mulder, M. and Hill, K. (2011). 'Controversies in the evolutionary social sciences: a guide for the perplexed', *Trends in Ecology and Evolution*, vol. 16(3), pp. 128–35.

Smith, K., Kalish, M.L., Griffiths, T.L. and Lewandowsky, S. (2008). 'Introduction: cultural transmission and the evolution of human behaviour', *Philosophical Transactions of the Royal Society, B*, vol. 363(1509), pp. 3469–76.

Stigler, G.J. and Becker, G.S. (1977). 'De gustibus non est disputandum', *American Economic Review*, vol. 67(2), pp. 76–90.

Trivers, R.L. (1971). 'The evolution of reciprocal altruism', *Quarterly Review of Biology*, vol. 46(1), pp. 35–57.

Williams, G.C. (1966). *Adaptation and Natural Selection*, Princeton: Princeton University Press.

Winterhalder, B. and Smith, E.A., editors (1992). *Evolutionary Ecology and Human Behavior*, New York: Aldine de Gruyter.

Review 2

1. Can Human Cooperation be Explained by Group Selection?

For some years, Samuel Bowles and Herbert Gintis have been working on an extraordinarily ambitious project that brings together economics, anthropology and biology. In this book, they summarise this work and present their latest thinking about it. As their title suggests, they propose the hypothesis that *homo sapiens* is a 'cooperative species'. By this they mean that, as a result of genetic natural selection, modern human beings are innately predisposed to cooperative behaviour. Bowles and Gintis propose that these cooperative predispositions evolved as adaptations to the natural and social environment in which ancestral humans lived before the development of agriculture. Their project is to try to identify the content of these predispositions and to explain the evolutionary processes by which they were selected.

The proposed explanation involves two broad forms of evolutionary theorising – *gene–culture co-evolution* and *multi-level selection*. The key idea of gene–culture co-evolution is that if a cultural practice – that is, a practice that is *not* genetically hard-wired – persists for a sufficiently long time in a sufficiently isolated population, it constitutes a niche in which genetic selection can operate. Thus, alongside the more normal mechanism by which cultural evolution within a population is constrained by that population's genetic endowment, gene pools can adapt to culture. Multi-level selection is the idea that natural selection can work, not only at the level of the individual but also at the level of groups of individuals. As it is uncontroversial that individual selection is a large part of natural selection, the key idea is that *group selection* played a significant role in the natural selection of human psychology.

When thinking about these evolutionary processes, it is essential to keep in mind the distinction between cultural and genetic evolution. The mechanisms of cultural and genetic evolution are isomorphic in important respects, and many of the tools of