



ECCE HOMO

Ladislav Kováč

Všetky objekty sveta, s ktorými my ľudia prichádzame do styku, sú buď prirodzené – vytvorené prírodou, alebo umelé, arteficiálne – výtvyry ľudskej kultúry ako procesu artefakcie. Preto sa zdá byť logické, ak vedy, ktoré skúmajú prírodné objekty, nazývame prírodnými vedami a vedy, ktoré skúmajú produkty kultúry, teda materiálne a symbolické artefakty, označujeme ako kultúrne vedy. Poznatky, ktoré ľudstvu priniesli prírodné vedy, a zároveň ich aplikácia v technike a ekonomike, dávajú našej dobe jej ohromujúcu dynamiku: dianie sa neustále zrýchľuje a jeho komplexnosť narastá do podoby, keď mu prestávame rozumieť a potom ho aj ovplyvňovať. Metódy prírodných vied sa stále zdokonaľujú, umožňujú nám siahať na jednej strane do hĺbok atómu a na druhej strane dovidieť až na priestorové a časové hranice vesmíru.

Úspechy v prienikoch k podstate hmoty a podstate vesmíru, ktoré dosiahli prírodné vedy, nemajú však svoju paralelu v úspechu v prieniku k podstate človeka, o ktoré sa nezávisle usilujú prírodné i kultúrne vedy. Otázka, ktorú si premýšľajúci ľudia kládli už od čias antiky, o tom, či je človek dobrý alebo zlý, anjel alebo diabol, nás dosiaľ stále znepokojuje. Ba dlho zostávala nezodpovedaná aj základná otázka „Čo je človek?“ Pred 40 rokmi, v roku 1976, sa biológ Richard Dawkins odvolal na zoológa Georgea Simpsona, ktorý napísal: „To, čo chcem teraz vysloviť je, že všetky snahy zodpovedať túto otázku pred rokom 1859 sú bezcenné a najlepšie urobíme, ak na ne celkom zabudneme.“ V zmienenom roku 1859 vyšla kniha Charlesa Darwina, v ktorej predložil svoju teóriu evolúcie života prirodzeným výberom. V roku 1986 Dawkins explicitne zdôraznil, že „naša vlastná existencia, ktorá kedysi predstavovala najväčšiu zo všetkých záhad, nie je viac záhadou, lebo bola rozriešená. Darwin a Wallace ju rozriešili a nám neostáva nič iné, než ešte po nejakú dobu písať poznámky pod čiarou k ich rozriešeniu“.¹

Vo svojej koncepcii „sebeckého génu“ Dawkins dokazoval, že človek je sebecký, egocentrický, sleduje iba svoj vlastný záujem, je teda od prírody zlý, neschopný altruizmu – obetovať svoj záujem v prospech iných ľudí. Paralelne s Dawkinsom iný prírodovedec, Edward O. Wilson, sa pokúsil problém sociálneho správania živočíchov – vrátane človeka – vykladať cez po-

¹ Richard Dawkins: *The blind watchmaker*. Northon, New York, 1987, s. ix.

dobnú koncepciu a na nej postaviť novú vedu, sociobiológiu, ktorá mala byť akýmsi mostom medzi prírodnými a kultúrnymi vedami. Koncepcia prírodovedca Dawkinsa bola aj v zhode s teóriou „racionálneho človeka“, ktorá slúžila jeho súčasníkom ekonomom pre výklad ľudského ekonomického správania. Ja sám som sa na stránkach *K & K* pripojil k tým, ktorí sa čitateľom snažili priblížiť učenie Dawkinsa.²

Dnes je jasné, že elegantná, ale simplexná teória „sebeckého génu“ bola mylná, práve tak ako teória „racionálneho človeka“ – a publikácia Gintisa a jeho kolegov, ktorú som teraz pre *K & K* odporučil, je jednou z tých, ktoré to vysvetľujú. (O tom, že zapôsobila aj na iných vedeckých kolegov, svedčí fakt, že bola dosiaľ až 999-krát citovaná.) Obidve teórie boli nielen mylné z hľadiska poznania, ale morálne zavádzajúce a škodlivé. Ostatne, sám tvorca sociobiológie, Edward O. Wilson, svoj výklad ľudského správania postupne skorigoval a prišiel s novou koncepciou ľudského správania, podľa ktorej ľudského jedinca charakterizujú dva protikladné sklony v ňom samom: k „dobru“ a „zlu“, takže je schopný konať cnosti rovnako ako páchať nerestí. Tým len tento prírodovedec konceptualizuje dnes to, o čom nás od nepamäti presviedčalo umenie svojimi špecifickými prostriedkami: že človek je protikladný živočích, ktorému chýba vnútorná konzistencia. Ale pekne to vyjadril aj filozof Michel Montaigne v 16. storočí vo svojich Esejách: „Sme zlátanina, tak beztvárá a zložením rôznorodá, že každý kúsok hrá v každej chvíli svoju vlastnú hru...“

Nová teória človeka, ktorá by mala byť výsledkom syntézy prírodných a kultúrnych vied, bude však ďaleko radikálnejšia, než je pohľad, ktorý v odporúčenom texte predkladajú behaviorálny vedec Gintis, antropológ Boyd a ekonómia Bowles a Fehr. Musí brať do úvahy aj lekciu, čo ľudstvo dostalo v 20. storočí, označovanom za „najkrvavejšie storočie dejín“, i lekciu z prvých dvoch dekád 21. storočia. Treba prehodnotiť dve základné idey, na ktorých bolo založené myslenie západnej kultúry: ideu racionalizmu a ideu individualizmu. Hoci hlavní myslitelia Západu stavali svoje teórie na viere v racionalizmus, výnimkou medzi nimi bol filozof 18. storočia David Hume: podľa neho človek nie je racionálny živočích; jeho správanie riadi nie rozum, ale emócie – rozum je len sluhom emócií. Lekcie, čo sme dostali, to potvrdzujú.

Na evolučnom konštruovaní človeka mali podiel ako evolúcia biologická, tak kultúrna, ale najmä aj samostatná evolúcia emocionálna. Emocionálna evolúcia pozemského života vyvrcholila na úrovni človeka: človek sa líši od ostatných živých tvorov v tom, že emócie nielen riadia jeho správanie, ale svoje emócie si aj uvedomuje, a to v podobe citov. Únik pred pocitom bolesti a honba za pocitom príjemnosti sa stali hlavnou motivačnou silou konania ľudského jedinca. Máme mozog zhotovený tak, aby každý z nás precítoval bolesti a príjemnosti nie iba svoje vlastné, ale aj iných ľudí: empatia k iným nás robí altruistickými. Iní ľudia sú tiež hlavným zdrojom našej osobnej príjemnosti. Takže človek nie je egocentrický, ale alterocentrický: v jeho myšli sú neustále prítomní iní ľudia. Zdieľaná príjemnosť sa znásobuje, zdieľaná bolesť znižuje.

Aby ľudský jedinec mohol prežiť, urobila z neho biologická evolúcia živočicha sociálneho: pri hľadaní potravy, obrane pred dravcami, starostlivosti o potomstvo museli jedinci spolupracovať. Ale emocionálna evolúcia zmenila človeka na tvora až hypersociálneho – človek je tak silno sociálny, ako sú mravce alebo včely. Ale k vedomému prežívaniu emócií v podobe citov dochádza iba na úrovni individuálneho človeka. Možno preto sa zrodil a postupne hypertrofoval v západnej kultúre individualizmus: presved-

2 Ladislav Kováč: „Človek je fanatik zmyslu“. In: *Kritika & Kontext*, 1/2000, s. 98 – 99.

čenie o suverenite ľudského jednotlivca a o sile individuálneho rozumu. Pozabudlo sa na sociálne emócie, ktoré silne rozšírili citový repertoár ľudského jedinca: pocity viny, hanby, závidi, nenávidi, ale aj lásky. Ony nám aj dokazujú, že každý z nás je súčasťou vyššej evolučnej entity, kedysi malej skupiny lovcov a zberačiek, neskôr národa, dnes aj nadnárodných útvarov, ale aj ľudského druhu ako celku. Sociálne emócie jednotlivca slúžia tejto vyššej entite – a z hľadiska jej záujmov je jeho prirodzenosť „dobrá“.

Hlavný dôkaz nám ale poskytuje kultúra. Biologická evolúcia vybavila človeka ako jediného živočícha rukou, a tým mu dala možnosť vytvárať artefakty, ktoré spustili kultúrnu evolúciu. Ale až emocionálna evolúcia pridala artefakcii jej jedinečnú podobu: artefakcia sa stala tvorbou a veľké potešenie z tvorby priviedlo na svet mýty, filozofiu, umenie a vedu. Artefakt, ktorý tvorca vytvorí, neslúži len jemu; slúži iným a tie najdokonalejšie z nich slúžia celému ľudstvu. Artefakcia je hlavnou formou ľudského altruizmu.

A čo tie lekcie z „najkrvavejšieho storočia histórie“ i zo záverečného 21. storočia? Presvedčajú nás o neprekročiteľných obmedzeniach, čo na nás uložila biologická evolúcia. Dokazujú nám, že človek je naozaj tvor nekonzistentný, so skoro neobmedzenými potenciami rovnako pre dobro ako pre zlo. Záleží od konkrétnej kultúry a od našich emócií, ktoré z týchto potencií sa presadia. Zlo jedinca je banálnym, ako konštatovala Hannah Arendtová pri analýze nacizmu. Ale dôsledná analýza zla nacizmu a rovnako aj komunizmu nás presvedča, že zlo inštitúcií, teda jedného zo sociálnych artefaktov, môže ľahko získať nadľudské rozmery a stať sa monštruóznym. Áno, Simpson mal pravdu: Darwin predstavuje rozhodujúci zlom v dejinách ľudského myslenia; tento zlom rozdelil tieto dejiny na dva obdobia: p. D. a po D. (pred a po Darwinovi). Prírodné vedy to vedia, ale malo by sa to stať aj samozrejmosťou v kultúrnych vedách.

Lenže múdrosť, čo sa nahromadila v ére p. D., nemožno ignorovať. Ukážkou je teória človeka, ktorú sformuloval Platón už pred 2400 rokmi a ktorá je aj dnes v súhlase s tým, k čomu dospela moderná veda: „Človek je, ako hovoríme, tvor krotký a ak sa mu dostane správnej výchovy a šťastnej prirodzenosti, stáva sa spravidla najbožskejším a najkrotkejším, ale keď sa vychováva nedostatočne alebo nedobre, najdivokejším, aké zem plodí.“ – Ale neplatí to len o výchove. Na formovaní podoby súčasného človeka sa podieľa spolu s prírodou, ktorej príkazy máme zabudované do našich génov, aj kultúra s poslaním ako ho videl už pred 350 rokmi Jan Amos Komenský: „šľachtiteľ ľudskú prirodzenosť“. Podstata človeka je neoddeliteľne práve tak kultúrna ako prírodná.

ECCE HOMO

Ladislav Kováč

All the objects in the world with which we humans come into contact are either natural products or artificial creations of human culture from the process of artefaction. Therefore, it appears to be logical for the sciences that research natural objects to be called the natural sciences and for the sciences that research the products of culture, whether material or symbolic artefacts, to be called cultural sciences. The findings humanity has obtained from the natural sciences and their application in technology and the economy have given our age its immense dynamism. Development is constantly accelerating and growing into a form in which we are ceasing to understand or influence it. The methods of the natural sciences are constantly being perfected. They enable us to reach in one direction into the depths of the atom and in the other to the spatial and temporal limits of the universe.

However, the successes in penetrating into the essence of matter and the essence of the universe achieved by the natural sciences, have no parallel in successes in penetrating into the essence of the human being for which the natural and cultural sciences are striving independently. The question thinking people have posed since antiquity, about whether the human being is good or evil, an angel or a devil, still disturbs us. The fundamental question: "What is a human being?" has long remained unanswered. Forty years ago, in 1976, the biologist Richard Dawkins referred to the zoologist George Simpson, who wrote, "What I want to say is that all the efforts to answer this question before the year 1859 are worthless and the best we can do is to forget about them." That year 1859 was when Charles Darwin published his book presenting his theory of the evolution of life by natural selection. In 1986, Dawkins explicitly emphasized that "Our own existence, which formerly represented the greatest of all mysteries, is no longer a mystery, because it has been solved. Darwin and Wallace solved it, and nothing remains to us, other than to write footnotes about their solution".¹

In his conception of the "selfish gene" Dawkins emphasized that human beings are selfish or egocentric, pursuing only their own interests. They are naturally evil and incapable of altruism or sacrificing their own interests in favour of other people. In parallel with Dawkins, another natural scientist, Edward O. Wil-



1 Dawkins, R.: The blind watchmaker. (Norton, New York, 1987, p. ix).

son, attempted to solve the problem of the social behaviour of animals, including humans through a similar conception and to build on it a new science, socio-biology, which aimed to be a sort of bridge between the natural and cultural sciences. The conception of the natural scientist Dawkins was also in harmony with the theory of the “rational person,” which served his economist contemporaries as a basis for explaining human economic behaviour. I have myself joined those who endeavoured to bring the views of Dawkins to the readers of *Kritika & Kontext*.²

Today, it is clear that the elegant, but simplistic theory of the “selfish gene” was erroneous, as was the theory of the “rational person.” The publication of Gintis and his colleagues, which I now recommend for *Kritika & Kontext*, is one of those that explain this. (Its influence on other scientific colleagues is shown by the fact that it has been cited 999 times.) Both theories were not only erroneous from the point of view of knowledge, but also morally misleading and harmful. Wilson, the creator of socio-biology, gradually corrected his interpretation of human behaviour, according to which human individuals have two opposing tendencies: towards “good” and towards “evil.” They are capable of acting virtuously and of pursuing vices. In this way, a natural scientist conceptualizes today the view expressed by art since time immemorial, namely that the human being is a contradictory creature lacking internal consistency. It was expressed very well in the sixteenth century by the philosopher Michel Montaigne in his *Essays*: “We are a patchwork, so shapeless and heterogeneous in composition that every part plays its own game at any moment.”

However, the new theory of the human being, which should be a result of the synthesis of natural and cultural sciences, will be much more radical than the view presented in the recommended text by the behavioural scientist Gintis, anthropologist Boyd, and economists Bowles and Fehr. It must also take into account the lesson humanity received in the twentieth century, described as the “bloodiest century in history,” and the lesson of the first two decades of the twenty-first century. It is necessary to re-evaluate two basic ideas on which the thought of Western culture was based: the idea of rationalism and the idea of individualism. Although the main thinkers of the West built their theories on faith in rationalism, an exception was the eighteenth-century philosopher David Hume. In his view, human beings are not rational creatures, and their behaviour is directed not by reason but by emotion—reason is only the servant of emotion. The lessons we have received confirm this.

Biological evolution, cultural evolution, and especially an independent emotional evolution shared in the evolutionary construction of the human being. The emotional evolution of life on Earth culminated on the human level. Human beings differ from other life forms in that emotions not only direct their behaviour, they are also aware of their emotions in the form of feelings. Escape from feelings of pain and pursuit of pleasant feelings became the main motivating forces for human activity. We have brains made so that each of us feels pain and pleasure not only for ourselves, but also for other people: empathy towards others makes us altruistic. Other people are also a major source of our personal pleasure. Therefore, a human being is not egocentric but alterocentric: other people are always present in his or her thoughts. Shared pleasure is multiplied, shared pain is reduced.

So that the human individual could survive, biological evolution made him or her into

2 Kováč, L.: Človek je fanatik zmyslu. (The human being is a fanatic for meaning) *Kritika & Kontext*, 2000, 1/00, pp. 98-99.

a social creature. Individuals had to cooperate in the search for food, defence against wild animals, and care for the young. However, emotional evolution changed the human being into a hyper-social creature. A human being is as strongly social as ants or bees, but the conscious experience of emotions in the form of feelings happens only on the level of the individual person. Perhaps this is why individualism was born and gradually hypertrophied in Western culture. It means belief in the sovereignty of the human individual and the strength of individual reason. The social emotions—feelings of guilt, shame, envy, hatred, but also love—which strongly extend the emotional repertoire of the human individual, were forgotten. They also prove to us that each of us is part of a higher evolutionary entity, formerly a small band of hunter-gatherers, later a nation, today also supra-national formations and the human species as a whole. The social emotions of the individual serve this higher entity—and from the point of view of its interests, its nature is “good.”

However, culture gives us the main proof of this. Biological evolution made the human being the only creature with hands, which provided the possibility to make artefacts, which launched cultural evolution. However, only emotional evolution gave artefact making its unique form: Artefact making became creation and great delight in creation led to the world of myth, philosophy, art, and science. An artefact a person creates does not serve only him or her, but also others, and the most perfect serve all humanity. Artefact making is the main form of human altruism.

What are the lessons from the “bloodiest century in history” and from the present twenty-first century? They convince us of the uncrossable limitations imposed on us by biological evolution. They prove to us that the human being is really an inconsistent creature with almost unlimited potential for good and evil equally. It depends on the specific culture and on our emotions, which apply these abilities. The evil of the individual is banal, as Hannah Arendt stated in her analysis of Nazism. But consistent analysis of Nazism and equally also of communism convinces us that the evil of institutions, of one of our social artefacts, can easily acquire superhuman dimensions and become monstrous. Yes, Simpson was right: Darwin represents a decisive turning point in the history of human thought. This turning point divides history into two parts: before Darwin and after Darwin. The natural sciences know this, but it should be just as obvious in the cultural sciences.

However, the wisdom accumulated in the era before Darwin cannot be ignored. An example is the theory of the human being formulated by Plato 2400 years ago and still in harmony with the views to which modern science has matured: “A human being is, as we say, a tame creature and if he or she gets the right upbringing and a fortunate nature, he or she usually becomes the most godlike and well-behaved, but when brought up inadequately or badly, he or she becomes the wildest thing the earth begets.” It does not apply only to upbringing. Culture also shares in the shaping of the modern person together with nature, the commandments of which are built into our genes. Culture has the mission of “improving human nature” as Jan Amos Comenius saw it 350 years ago. The essence of a human being is inseparably both cultural and natural.

EXPLAINING ALTRUISTIC BEHAVIOR IN HUMANS

Herbert Gintis,
Samuel Bowles, Robert Boyd,
and Ernst Fehr

INTRODUCTION

The explanatory power of inclusive fitness theory and reciprocal altruism (Hamilton, 1964; Trivers, 1971; Williams, 1966) convinced a generation of researchers that what appears to be altruism—personal sacrifice on behalf of others—is really just long-run self-interest. Richard Dawkins (1976, 1989), for instance, struck a responsive chord when, in *The Selfish Gene*, he confidently asserted “We are survival machines—robot vehicles blindly programmed to preserve the selfish molecules known as genes... This gene selfishness will usually give rise to selfishness in individual behavior.” Dawkins allows for morality in social life, but it must be socially imposed on a fundamentally selfish agent. “Let us try to teach generosity and altruism,” he advises, “because we are born selfish.” Yet, even social morality, according to R. D. Alexander, the most influential ethicist working in the William-Hamilton tradition, can only superficially transcend selfishness. In *The Biology of Moral Systems*, Alexander (1987, p. 3) asserts, “ethics, morality, human conduct, and the human psyche are to be understood only if societies are seen as collections of individuals seeking their own self-interest.” In a similar state of explanatory euphoria, Ghiselin (1974, p. 247) claims “No hint of genuine charity ameliorates our vision of society, once sentimentalism has been laid aside. What passes for cooperation turns out to be a mixture of opportunism and exploitation... Scratch an altruist, and watch a hypocrite bleed.”

However, recent experimental research has revealed forms of human behavior involving interaction among unrelated individuals that cannot be explained in terms of self-interest. One such trait, which we call strong reciprocity (Gintis, 2000b; Henrich et al., 2001), is a predisposition to cooperate with others and to punish those who violate the norms of cooperation, at personal cost, even when it is implausible to expect that these costs will be repaid either by others or at a later date.

In this article, we present empirical evidence supporting strong reciprocity as a schema for explaining important forms of altruism in humans. We then explain why, under conditions plausibly characteristic of the early stages of human evolution, a small fraction of strong reciprocators could invade a population of self-regarding types, and why strong reciprocity is an evolutionarily stable strategy. Although most of the evidence we report is based on behavioral experiments, the same behaviors are regularly observed in everyday life, for example in wage setting by firms (Bewley, 2000), tax compliance (Andreoni, Erard, & Feinstein, 1998), and cooperation in the protection of local environmental public goods (Acheson, 1988; Ostrom, 1990).

In supporting the importance of strong reciprocity, we of course do not deny the

importance of either kin altruism (Hamilton, 1964) or reciprocal altruism (Trivers, 1971). Both are beyond doubt potent forces in human motivation. We do believe, however, that the evolutionary success of our species and the moral sentiments that have led people to value freedom, equality, and representative government are predicated upon strong reciprocity and related motivations that go beyond inclusive fitness and reciprocal altruism.

We wish to avoid three common misunderstandings of our argument. First, many contemporary researchers reject our critique of Dawkins, Alexander, and others in the “selfish gene” school by asserting that their pronouncements should not be taken at face value. Rather, they say, references to phenotypic behavior as “selfish” should be understood as asserting that the underlying genetic structures are subject to Darwinian evolutionary forces. Yet, these authors understood that their assertions were likely to be taken at face value, rather than being dramatic circumlocutions expressing completely unexceptionable propositions. It is only plausible, then, to suggest that they meant them, that it was plausible at the time to make such statements, but that they are now seen to be incorrect.

Second, we are often interpreted as rejecting the “gene-centered” approach to modeling human behavior. In fact, our results in no way contradict the standard population biology approach to genetic and cultural change. A gene that promotes self-sacrifice on behalf of others will die out unless those helped carry the mutant gene or otherwise promote its spread. In a population without structured social interactions of agents, behaviors of the type found in our experiments and depicted in our models could not have evolved. However, multilevel selection and gene–culture coevolutionary models support cooperative behavior among nonkin (Bowles, Choi, & Hopfensitz, in press; Feldman, Cavalli-Sforza, & Peck, 1985; Gintis, 2000, in press-a, in press-b; Henrich & Boyd, 2001; Sober & Wilson, 1998). These models, some of which are discussed below, are not vulnerable to the classic critiques of group selection by Dawkins (1976), Maynard Smith (1976), Rogers (1990), Williams (1966), and others.

Third, we are often told that the behavior we describe can in fact be explained by standard individual selection, kin selection, and reciprocal altruism models applied to the ancestral natural and social environment to which our species was subject during the period of its evolutionary emergence, where anonymous, one-shot interactions were supposedly extremely rare. Strong reciprocity in contemporary environments, according to this view, is a maladaptation. We think this alternative is unlikely, and address the issues in Section 8.

7. THE COEVOLUTION OF INSTITUTIONS AND BEHAVIORS

If group selection is part of the explanation of the evolutionary success of cooperative individual behaviors, then it is likely that group-level characteristics—such as relatively small group size, limited migration, or frequent intergroup conflicts—that enhance group selection pressures coevolved with cooperative behaviors. Thus, group-level characteristics and individual behaviors may have synergistic effects. This being the case, cooperation is based in part on the distinctive capacities of humans to construct institutional environments that limit within-group competition and reduce phenotypic variation within groups, thus heightening the relative importance of between-group competition, and hence allowing individually costly but in-group-beneficial behaviors

to coevolve with these supporting environments through a process of interdemic group selection.

The idea that the suppression of within-group competition may be a strong influence on evolutionary dynamics has been widely recognized in eusocial insects and other species. Boehm (1982) and Eibl-Eibesfeldt (1982) first applied this reasoning to human evolution, exploring the role of culturally transmitted practices that reduce phenotypic variation within groups. Examples of such practices are leveling institutions, such as monogamy and food sharing among nonkin, namely those that reduce within-group differences in reproductive fitness or material well-being. By reducing within-group differences in individual success, such structures may have attenuated within-group genetic or cultural selection operating against individually costly but group-beneficial practices, thus giving the groups adopting them advantages in intergroup contests. Group-level institutions thus are constructed environments capable of imparting distinctive direction and pace to the process of biological evolution and cultural change. Hence, the evolutionary success of social institutions that reduce phenotypic variation within groups may be explained by the fact that they retard selection pressures working against in-group-beneficial individual traits and the fact that high frequencies of bearers of these traits reduces the likelihood of group extinctions.

We have modeled an evolutionary dynamic along these lines, exploring the possibility that intergroup contests play a decisive role in group-level selection. Our models assume that genetically and culturally transmitted individual behaviors, as well as culturally transmitted group-level characteristics, are subject to selection (Bowles, 2001; Bowles, Choi, & Hopfensitz, 2002). We show that intergroup conflicts may explain the evolutionary success of both (a) altruistic forms of human sociality towards nonkin and (b) group-level institutional structures such as food sharing and monogamy that have emerged and diffused repeatedly in a wide variety of ecologies during the course of human history. In-group-beneficial behaviors may evolve if (a) they inflict sufficient costs on out-group individuals and (b) group-level institutions limit the individual costs of these behaviors and thereby attenuate within-group selection against these behaviors.

Our simulations show that if group-level institutions implementing resource sharing or nonrandom pairing among group members are permitted to evolve, group-beneficial individual traits coevolve along with these institutions, even where the latter impose significant costs on the groups adopting them. These results hold for specifications in which cooperative individual behaviors and social institutions are initially absent in the population. In the absence of these group-level institutions, however, group-beneficial traits evolve only when intergroup conflicts are very frequent, groups are small, and migration rates are low. Thus, the evolutionary success of cooperative behaviors in the relevant environments during the first 90,000 years of anatomically modern human existence may have been a consequence of distinctive human capacities in social institution-building (Boyd & Richerson, in press).

CONCLUSION

Much more experimental and theoretical work must be done to understand the major outlines of human prosocial behavior. We suspect, on the basis of the many studies completed over the past several years, that the new knowledge obtained will give us

a picture of prosociality (and its obverse, antisociality) that is fundamentally incompatible with the economist's model of the self-interested actor and the biologists' model of the self-regarding reciprocal altruist.

Contemporary behavioral theory is the legacy of several major contributions (Cosmides & Tooby, 1992; Dawkins, 1989; Hamilton, 1964; Maynard Smith, 1982; Trivers, 1971; Williams, 1966; Wilson, 1975), all of which assumed the relations between nonkin could be modeled using self-interested actors. It is not surprising, then, that the most successful research in behavioral theory has been in the area of the family, kinship, and sexual relations, while the attempts to deal with the more complex interactions characteristics of social group behavior have been less persuasive. To address this situation, we believe that more attention should be paid to (a) the origin and nature of social emotions (including guilt, shame, empathy, ethnic identity, and ethnic hatred), (b) to the coevolution of genes and culture in human social history, (c) the role of group structure and group conflict in human evolution, and (d) integrating sociobiological insights into mainstream social sciences.

Fragments from 9-a page article in *Evolution and Human Behavior*, no. 24 (2003), pp. 153 - 172.

The whole text can be found here: <https://www.econ.uzh.ch/dam/jcr:fffff-9758-127f-ffff-ffffae2365e4/ExplainingAltruisticBehaviorinHumans.pdf>