ELSEVIER

Journal of Economic Behavior & Organization
Vol. 43 (2000) 91–99

JOURNAL OF
Economic Behavior
& Organization

www.elsevier.com/locate/econbase

On some implications of evolutionary psychology for the study of preferences and institutions☆

Avner Ben-Ner[a], Louis Putterman[b,*]

[a] *Industrial Relations Center, University of Minnesota, Minneapolis, MN 55455, USA*
[b] *Department of Economics, Brown University, Providence, RI 02912, USA*

Received 8 April 1997; received in revised form 19 January 2000; accepted 9 February 2000

Abstract

In many economic interactions, for instance in firms, the standard approximation of strict self-interest is inadequate to modeling human behavior. A scientific theory of preferences, grounded in evolutionary psychological and biological theory, can avoid resort to ad hoc assumptions. Evolutionary theory is supported by a growing body of data including new results in experimental economics. It holds that the evolved human nature includes an ability to solve social dilemma problems through reciprocity and punishment of cheaters. Treating realized preferences as phenotypic expressions with both environmental and genetic causes will also allow economists to study the impact of institutions on preferences. © 2000 Elsevier Science B.V. All rights reserved.

JEL classification: D00; C91; C72

Keywords: Evolution; Preferences; Evolutionary psychology; Altruism; Reciprocity

1. Introduction

Darwinism was once taken to imply that the process of biological competition can spare no sentimentality for unselfishness or indeed, for sentiment itself. The most ruthless, self-seeking, and rational organisms will survive, while weaker members of their species will fall by the wayside. Darwinism and the economics of the century that followed *The Origin of Species* appeared to entail kindred social theories, in which human moral pretensions were set aside and the natural competitiveness of the species claimed its full due.

☆ We would like to thank Herbert Gintis, Robert Wright, and three anonymous referees and the editor for helpful comments on earlier drafts.

* Corresponding author. Tel.: +1-401-863-3836; fax: +1-401-863-1970.

PII: S0167-2681(00)00110-4

Edgeworth's dictum that "The first principle of economics is that every agent is actuated only by self-interest" [1] was thought by many to be Darwin's as well.

Darwin (1871) himself had been interested in both sentiment and altruism. But while this fact was known to some careful readers, it took many decades before the cooperative dimension of social interaction emerged as a central focus of biology and allied social sciences. This new focus is exemplified by the work of biologist Robert Trivers in the 1970s, political scientist Robert Axelrod and biologist William Hamilton in the 1980s, and psychologist Leda Cosmides and anthropologist John Tooby in the 1990s, although other important contributors could be cited. [2] Theorists of sociobiology and evolutionary psychology take as their starting point the idea that it is genes, rather than the organisms that bear them, that engage in the struggle for survival. From an evolutionary standpoint, survival of several replicas is better than survival of any one copy. Organisms not predisposed by their genes to invest resources in bearing and (where immaturity at birth requires it) nurturing offspring will make poor agents for "replication-minded" genes. Any undue propensity to shirk on investments in reproduction and nurturing, should it appear by the serendipity of mutation, will be selected against, since the genes in which the new propensity is encoded will leave fewer copies in future generations. More broadly, propensities to invest resources in helping close kin, who share a significant fraction of one's own genes, will pay off in evolutionary terms. [3]

1.1. Altruism and reciprocity: not all is self-interest

The universality of maternal and, more broadly, of kin altruism across human cultures is an illustration of the fact that organic evolution can produce organisms that are not strictly self-interested. But more remarkable and of at least equal importance to economics and other social sciences is the phenomenon of reciprocal cooperation among *nonkin*. Sympathies among nonkin are arguably rare or shallow, or where they exist, they might be supposed to arise through a "misplaced" diffusion of kin altruism. Yet in the small band-like societies that are believed to have constituted the environment of most of human evolution, propensities to cooperate with other group members and to react with anger toward nonreciprocators of cooperation may have been genetically favored. Large-brained hominids who repeatedly interacted would have learned to recognize one another's behaviors and to seek out others known to behave cooperatively when engaging in interactions with positive sum payoffs (for instance, sharing meat from hunting excursions character-

[1] Edgeworth (1881).

[2] See Trivers (1971), Axelrod and Hamilton (1981). In this note, we concentrate especially on the field called evolutionary psychology. See Barkow et al. (1992), Buss (1995), and for an accessible popular treatment and further references, see Wright (1994). Evolutionary psychology forms part of a broader set of evolutionary approaches that includes the well-known work of Dawkins (1989), the sociobiological approach of Wilson (1978), and the coevolutionary approaches of Cavalli-Sforza and Feldman (1981), Lumsden and Wilson (1983), Boyd and Richerson (1999), and Durham (1991). We also mention in a later note but do not discuss at length the group selectionist variant represented by Sober and Wilson (1998). Finally, we do not attempt to review exhaustively past discussions of sociobiology by economists; for a wide-ranging treatment, see Hirshleifer (1977).

[3] The mathematics of selection for kin altruism was first developed by biologist Hamilton (1964); a discussion with applications for economics is provided by Bergstrom (1996, pp. 1905–1908).

ized by high variance of amounts procured). The presence within a human population of feelings of anger against noncooperators, fueling acts of costly punishment that appear irrational from the standpoint of individual interest, would have helped to deter cheating and to raise the evolutionary payoff of bearing the genetic inclination to cooperate. Increases in: (a) the strength of the inclination to cooperate, (b) the cognitive capacity to recognize cooperators, detect cheaters, and remember who was who, and (c) the inclination to punish noncooperators, could therefore have been mutually reinforcing evolutionary trends. [4]

While the possibility of cooperation in a repeated prisoners' dilemma of unknown duration also exists for the rational egoist, what Cosmides and Tooby (1992) call "cognitive adaptations for social exchange" — including a heightened ability to detect cheating — differ from the ability to undertake unconstrained optimization that is assumed in standard economic models. In the Cosmides–Tooby model, the individual is endowed with cognitive predispositions that generate patterns of behavior under the influence of environmental stimulae. Selection requires that the behavioral patterns generated by a predisposition be optimal for the genes, rather than that choices made in particular instances be optimal for the gene-bearer. Evolution does not bequeath to the gene-bearer a set of *general* goals, leaving that individual to *reason* out how to implement them. Rather, one is given certain emotive and cognitive predispositions, the purposes (and indeed the presence) of which one need not fathom. [5]

Evolutionary psychology is not hostile to the view that people pursue self-interest. Traditional economics got by with the assumption of selfish individuals because that assumption is a good approximation, a good deal of the time. When individuals interact more or less anonymously and in relatively large groups, the assumption that their actions are mainly geared to promoting the interests of themselves and their immediate kin is often a sensible one. Even in the large group, anonymous setting, the assumption can go awry, however. Thus, cooperation in large one-shot prisoners' dilemma or voluntary contribution mechanism experiments is typically about 50 percent, and contribution rates are actually higher

[4] To what degree the selection had to have operated at the *group* level in order for these traits to become established is a controversial question. The individual and her immediate kin may have been able to gain an advantage by free-riding on reciprocity and on the costly punishment of cheating on reciprocity provided by *other* group members. If such behaviors significantly reduced group outcomes and if mobility across groups was limited, however, the fates of close kin and of the groups to which they belonged would have been closely linked. On the role of group selection, see inter alia Sober and Wilson (1998), and Boyd and Richerson (1999). A detailed discussion of group selection is beyond the scope of this note.

[5] Thus, evolution builds behavioral predispositions using modules of discrete behavioral propensities, rather than flexible all-purpose objective functions. Another important illustration relates to the classic problem of how contemporary human beings could carry traits genetically selected to maximize reproductive success, and yet intentionally choose to limit the number of their offspring. The answer of evolutionary psychology is that the mechanisms used by evolution to foster our reproduction in paleolithic times, principally the sex drive and nurturative inclinations towards offspring, need not imply any conscious or unconscious drive towards self-replication per se. The fact that our pre-historically given traits may not be well adapted (from our genes' point of view) to modern opportunities like the discovery of birth control devices illustrates the proposition that we are adapted to past, but not necessarily to present, environments. Note that in this example as in others in which the interests of "replication minded genes" and of human individuals may not coincide, we are compelled neither by reason nor by moral principles to accept the gene's view as a morally superior one.

in larger groups (Isaac and Walker, 1988). People take the trouble to vote and to recycle reusable materials. [6]

1.2. The capacity for nastiness

It is by no means the case that all of the non-neoclassical implications of the evolutionary approach are "nice" ones like inclinations to respond to kindness with kindness. This is already clear from the fact that the impulse to punish noncooperation is a central element of reciprocal fairness. While arguably critical to sustaining cooperation, the same tendency to retaliate sometimes leads to tragic vendettas, blood feuds, and ethnic conflicts. Furthermore, the individuals predicted by evolutionary psychology are highly attuned to differences in inclusive-fitness interests that arise even among close relatives. Conflicts in the parent–child relationship, for instance, are to be expected because the genetically-determined preferred balance between the additional investment in a given offspring and investment in producing additional offspring will sooner or later diverge between the mother and child, with the latter caring more on the margin than does the former about its own survival and future reproduction, as opposed to the birth, survival and reproduction of siblings. Human beings evolve to be sensitive, as parents, to their child's cries, as serves their genetic interest; but it is in the interest of the child's genes to exploit this sensitivity at the occasional expense of sibling and parent alike. The child's inherited receptivity to parental messages regarding altruism towards the parents and siblings is in its turn exploited by the parent to attenuate these conflicts.

"Nastiness" is also thought to be endemic in relation among the sexes, as the asymmetry of physiological limits on the number of offsprings would lead males to develop capacities for feigning commitment to females. The latter become astute readers of the signs of such commitment, since paternal investment may crucially affect the survival of the relatively few offspring they can bear. In the males' genetic interests, the feigning can reach the point of successful self-deception; yet in that same interest, the male is inclined to seek opportunities to parent offspring with other females. In a similar vein, in contexts that support that reciprocal cooperation with nonkin to which reference was made above, it is in the genetic interest of the individual *to be perceived to be* a cooperator, not necessarily to cooperate. Where the perception can be safely cultivated while conserving on actual resources invested, genetic interest would dictate that this be done. There is also a prediction of self-deception: to fully succeed in evincing cooperativeness, we convince ourselves that we are indeed trustworthy, loyal, and moral, all the while on the look-out for opportunities to shade on reciprocity towards nonkin in favor of our immediate genetic interest. [7]

These remarks on deception and self-deception lead us to shift for a moment from the question of human selfishness to that of human rationality and freedom from sentiment. Evolutionary psychologists suggest that human beings have been adapted for rational information processing and objective self-awareness only to the extent that these qualities

[6] We thank a referee for the examples.

[7] Frank's (1988) argument that it is often better to develop genuine commitment to "niceness" can be viewed as an interesting refinement on this issue that deserves testing against the variant presented by Wright (1994) and others.

improved inclusive fitness in ancestral environments. Where self-deception would provide a reproductive edge, it might be found in abundance. While the development of mental capacities far exceeding those of other species on our planet is an obvious hallmark of human beings, this does not mean that we are optimally designed all purpose computers. On the contrary, we come loaded with pre-adapted cognitive machinery that evolved over millions of years (Barkow et al., 1992). Given the cumulative character of biological change — that is, its tendency to build opportunistically on pre-existing structures — the evolutionary approach leads us to expect that the mind is most unlikely to be a perfect reasoning machine. [8] The old idea that cognitive processes are most efficient when shorn of emotion is giving way to a new view, according to which the brain's emotional processes play a central role in guiding the thinking mind towards relevant solutions and away from unending rational loops. [9]

1.3. Endogenous preferences and institutions

In a view that we develop at greater length in Ben-Ner and Putterman (1998), the conjunction of: (a) a postulated genetic basis for human behavioral predispositions, and (b) the demonstrable impact of environment on phenotypic variation in behavior, opens up the possibility of a scientific research program for studying the influence of human environments on human preferences. The research program we envision is one that endogenizes preferences to economic and social environments. There exist a number of important economic and social problems, including those of cooperation and loyalty in organizations, the provision of nurturing and supportive environments for children and the elderly, the viability of humane social insurance mechanisms, responses to addiction and criminality, and even the creation and preservation of individuals who support the ethical norms sustaining low-cost exchange relations, with respect to which the economic method of analysis involving individual optimization subject to constraints can be more usefully applied if scope is allowed for extended preferences (Aaron, 1994; Bowles, 1998; Rabin, 1998). Understanding of such issue calls for a realization of the mutual impacts of institutions and environments upon preferences, including normative ones, and conversely the influence of such preferences on the performance, viability, and ultimately the very selection of such institutions and environments. The successful avoidance of intellectual chaos as we explore the largely uncharted territory of preference endogeneity, however, may depend upon the availability of a guiding paradigm regarding the nature and sources of human behavioral predispositions.

The evolutionary approach is rich not only in its allowance for an "extended" model of preferences, but also in its recognition that realized preferences are the result of *both* inherited receptivities *and* of the way in which experiences of the individual impact upon those inheritances. The biologists' distinction between a genotype, the sum of the genetic instructions provided to an organism, and a phenotype, the realized organism dependent

[8] See the accessible introduction to the architecture of the brain in Goleman (1995), the arguments on reason and emotion by Damasio (1994) and Elster (1998), and for an excellent recent discussion of cognition and rationality in economics, with further references, see MacLeod (1996).

[9] See again Damasio (1994).

on the interaction of those instructions with a particular environment, is especially useful for the study of human preferences. In certain insect species (for instance locusts), the individual is prompted by environmental stimuli (in that example, the degree of crowding with other members of the species under which it matures) to become one or another of two or more radically different types of organisms marked by differing physiology and behavior despite possession of an identical set of genes. [10] In a similar manner, human beings are influenced by the environments of their upbringing and socialization, as well as those in which they live as adults, to develop one or another set of preferences, including ones traditionally associated with moral values. Thus, much as recent cognitive research has shown that the development of specific areas of the brain is influenced by an individual's exposure to relevant environmental stimulae during a critical period of growth, so future research may show how exposure to different normative signals in formative periods leads to differing actualizations of the social potentials latent in our genes.

Both the importance of environment and the facts of sexual reproduction mean that evolutionary psychology does not predict perfect homogeneity of human beings. Tooby and Cosmides (1990) theorize that sexual reproduction, which induces variation in the protein structures of our cells, emerged as a strategy for thwarting the successful destruction of the complex, large-bodied host species by rapidly multiplying and evolving parasites. While basically alike except in instances of gross genetic error, humans differ in numerous details, including the settings of mechanisms regulating fear and other emotions, and thus *conceivably* also in receptiveness towards moral exhortation. Genetic differences in receptiveness to particular cultural messages are probably randomly distributed across populations: the genetic make-up of an average individual in Hungary or Peru differs little from that of an average individual in India or Japan (Tooby and Cosmides, 1990), so children with Hungarian genes can be as easily socialized to be culturally Peruvian, Indian, or Japanese, depending on how they are raised. Yet variations across individuals within any given society, and variation of environmental stimuli both within and across societies, lead to substantial differences in behavioral inclinations of specific human beings. This may explain the heterogeneity of responses to the experimental game situations both within and between populations in distinct cultural environments (Roth et al., 1991). [11]

Phrases such as "genetic predisposition," which recur frequently in the evolutionary psychology literature, imply behavioral inclinations that are literally traceable to genes or to gene–environment interactions. While *direct* evidence of genetic impacts on behavior is still limited for humans, [12] the body of *indirect* evidence that bears on the hypotheses about human nature derived from evolutionary psychology is large and growing. In addition to the ethnographic, sociological, and psychological observations that provide support for the evolutionary approach, work by experimental economists that has bearing upon the predictions of conditional altruism, reciprocity, moralistic aggression, and imperfect rationality is

[10] One form of the locust has shorter wings, longer legs, and a narrower pronotum with higher crest and larger head and is comparatively quiet and solitary, while the other form has a more saddle-shaped pronotum, broader shoulders and long wings and is comparatively active, nervous, and gregarious.

[11] Behavioral differences even sharper than those among the cultural groups studied by Roth et al. were found in recent experimentation with subjects drawn from an indigenous group inhabiting the southeastern Peruvian Amazon, the Machiguenga (Henrich, 2000).

[12] See, however, the summary of recent evidence on gene-behavior linkages by Hamer and Copeland (1998).

now beginning to accumulate rapidly. For instance, the tendency to engage in "irrational" costly punitive acts is remarkably illustrated by the experiments of Fehr and Gächter (2000). They demonstrate that introducing the option to penalize into a voluntary contribution game changes the convergence towards static Nash equilibrium that characterizes the standard experiment into high and rising levels of cooperation, even though punishing is never a dominant strategy in their experimental set-up. Kahneman et al. (1986), Roth et al. (1991), Fehr et al. (1993), Frank et al. (1993), Camerer and Thaler (1995), Hoffman et al. (1996, 1998), McCabe et al. (1996), Fehr and Gächter (1998), and Ben-Ner et al. (1999) provide other examples of experimental findings consistent with evolutionary assumptions about preferences.

1.4. Towards a scientific view of human nature

Economists are often heard to complain that the attempt to explain some behavior which appears to defy standard neoclassical theory by appeal to additional arguments in the utility function are ad hoc in character. The goal of economics, to such practitioners, is to take the assumptions of strict self-interest and rationality and to explain as much of actual behavior as possible with them. To allow for extendedness of preferences is viewed, from this standpoint, as an instance of "cheating by changing the rules of the game." While this game could well go on for many years into the future, we think that in the long-run, the economics discipline as a whole will recognize that the old assumption of rational, strictly self-interested individuals is not only an inexact and special approximation, but also inconsistent with a scientific view of human nature as the product of an evolutionary process. Models of rational, self-interested behavior are likely to persist and be put to many uses, but they will be recognized as special cases. What the evolutionary sciences are providing is the foundation for a nonarbitrary reformulation of individual choice theory along scientific lines. As psychologist Donald Campbell put it a decade ago, "Rationality in economic theory is primarily a rationality of means . . . While it is a maximization of *self-interest*, the contents of self-interest are usually not specified by the theory. Merging such self-interest theory with evolutionary biology offers the promise of theoretical grounds for predicting such contents, that is, predicting what sorts of interest the products of biological evolution would be apt to have." [13]

References

Aaron, H., 1994. Public policy, values, and consciousness. Journal of Economic Perspectives 8, 3–21.

Axelrod, R., Hamilton, W., 1981. The evolution of cooperation. Science 211, 1390–1396.

Barkow, J., Cosmides, L., Tooby, J. (Eds.), 1992. The Adapted Mind: Evolutionary Psychology and the Generation of Culture. Oxford University Press, New York.

Ben-Ner, A., Putterman, L., 1998. Values and institutions in economic analysis. In: Ben-Ner, A., Putterman, L. (Eds.), Economics, Values and Organization. Cambridge University Press, New York, pp. 3–69.

[13] Campbell (1989, p. 172). Economist Theodore Bergstrom articulates a similar position when he writes: "An evolutionary perspective . . . is likely to produce deeper insights and better-posed questions than theory based on arbitrary assumptions about preferences." See Bergstrom (1996, p. 1930).

Ben-Ner, A., Putterman, L., Magan, D., Kong, F., 1999. Reciprocity in a two part dictator game. Working Paper No. 99-28. Department of Economics, Brown University.

Bergstrom, T., 1996. Economics in a family way. Journal of Economic Literature 34, 1903–1934.

Bowles, S., 1998. Endogenous preferences: the cultural consequences of markets and other economic institutions. Journal of Economic Literature 36, 75–111.

Boyd, R., Richerson, P., 1999. Population structure, equilibrium selection and the evolution of group beneficial norms. Paper presented at the Meeting of the Allied Social Science Associations, New York, January 1999.

Buss, D., 1995. Evolutionary psychology: a new paradigm for psychological science. Psychological Inquiry 6, 1–30.

Camerer, C., Thaler, R.H., 1995. Ultimatums dictators and manners. Journal of Economic Perspectives 9, 209–219.

Campbell, D., 1989. Rationality and utility from the standpoint of evolutionary biology. In: Hogarth, R., Reder, M. (Eds.), Rational Choice: The Contrast Between Economics and Psychology. University of Chicago Press, Chicago, pp. 171–180.

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

Cosmides, L., Tooby, J., 1992. Cognitive adaptations for social exchange. In: Barkow, J., Cosmides, L., Tooby, J. (Eds.), The Adapted Mind: Evolutionary Psychology and the Generation of Culture. Oxford University Press, New York, pp. 163–228.

Damasio, A., 1994. Descartes' Error: Emotion, Reason, and the Human Brain. Putnam, New York.

Darwin, C., 1871. The Descent of Man, and Selection in Relation to Sex. Appleton, New York.

Dawkins, R., 1989. The Selfish Gene, 2nd Edition. Oxford University Press, Oxford.

Durham, W., 1991. Coevolution: Genes, Culture, and Human Diversity. Stanford University Press, Stanford.

Edgeworth, F.Y., 1881. Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences. Paul, London.

Elster, J., 1998. Emotions and economic theory. Journal of Economic Literature 36, 47–74.

Fehr, E., Gächter, S., 1998. How effective are trust- and reciprocity-based incentives? In: Ben-Ner, A., Putterman, L. (Eds.), Economics, Values and Organization. Cambridge University Press, New York.

Fehr, E., Gächter, S., 2000. Cooperation and punishment. American Economic Review, in press.

Fehr, E., Kirchsteiger, G., Riedl, A., 1993. Does fairness prevent market clearing? An experimental investigation. Quarterly Journal of Economics 108, 437–460.

Frank, R., 1988. Passions Within Reason: The Strategic Role of the Emotions. Norton, New York.

Frank, R., Gilovich, T., Regan, D., 1993. Does studying economics inhibit cooperation? Journal of Economic Perspectives 7, 159–171.

Goleman, D., 1995. Emotional Intelligence. Bantham Books, New York.

Hamer, D., Copeland, P., 1998. Living With our Genes: Why they Matter More than You Think. Doubleday, New York.

Hamilton, W., 1964. The genetical evolution of social behavior. Parts I and II. Journal of Theoretical Biology 7, 1–52.

Henrich, J., 2000. Does culture matter in Economic Behaviour? Ultimatum Game Bargining Amoung the Machiguenga of the Peruvian Amazon, American Economic Review, in press.

Hirshleifer, J., 1977. Economics from a biological viewpoint. Journal of Law and Economics 20, 1–52.

Hoffman, E., McCabe, K., Smith, V., 1996. Social distance and other-regarding behavior in dictator games. American Economic Review 86, 653–660.

Hoffman, E., McCabe, K., Smith, V., 1998. Behavioral foundations of reciprocity: experimental economics and evolutionary psychology. Economic Inquiry 36, 335–352.

Isaac, R.M., Walker, J.M., 1988. Group size effects in public goods provision: the voluntary contributions mechanism. Quarterly Journal of Economics 103, 179–200.

Kahneman, D., Knetsch, J., Thaler, R., 1986. Fairness as a constraint on profit seeking: entitlements in the market. American Economic Review 76, 728–741.

Lumsden, C., Wilson, E.O., 1983. Promethean Fire: Reflections on the Origin of Mind. Harvard University Press, Cambridge, MA.

MacLeod, B., 1996. Decision, contract and emotion: some economics for a complex and confusing world. Canadian Journal of Economics 29, 788–810.

McCabe, K., Rassenti, S., Smith, V., 1996. Game theory and reciprocity in some extensive form bargaining games. In: Proceedings of the National Academy of Science, November, pp. 13421–13428.
Rabin, M., 1998. Psychology and economics. Journal of Economic Literature 36, 11–46.
Roth, A., Prasnikar, V., Okuno-Fuhiwara, M., Zamir, S., 1991. Bargaining and market behavior in Jerusalem, Ljubljana, and Tokyo: an experimental study. American Economic Review 81, 1068–1095.
Sober, E., Wilson, D.S., 1998. Unto Others: The Evolution and Psychology of Unselfish Behavior. Harvard University Press, Cambridge, MA.
Tooby, J., Cosmides, L., 1990. On the universality of human nature and the uniqueness of the individual: the role of genetics and adaptation. Journal of Personality 58, 17–67.
Trivers, R.L., 1971. The evolution of reciprocal altruism. Quarterly Review of Biology 46, 35–57.
Wilson, E.O., 1978. On Human Nature. Harvard University Press, Cambridge, MA.
Wright, R., 1994. The Moral Animal: The New Science of Evolutionary Psychology. Vintage Books, New York.