

UTILITARIAN IMPARTIALITY AND CONTEMPORARY DARWINISM*

TOMISLAV BRACANOVIĆ, Zagreb, Chorvátsko

BRACANOVIĆ, T.: Utilitarian Impartiality and Contemporary Darwinism
FILOZOFIA 62, 2007, No 1, p. 14

One important component of utilitarian ethical framework is a specific version of the principle of impartiality. The principle claims that one should bring about the greatest possible overall utility or happiness and that our moral and morally relevant actions ought to result from objective and neutral deliberations, with all our personal interests, likes and dislikes left out. Drawing on relevant insights from sociobiology and evolutionary psychology this paper seeks to show that utilitarian principle of impartiality is seriously endangered by two interconnected Darwinian facts: (1) the fact that human biologically shaped psychology and motivation mechanisms do not fit well with the principle requiring the indiscriminate promotion of general happiness, and (2) the fact that human beings are individuals with naturally evolved personal desires, projects and ideals, and not merely spare parts of some general utilitarian clockwork of happiness.

Key Words: Utilitarianism – Impartiality – Darwinism

1. Utilitarianism and the principle of impartiality. One important and practically indispensable component of almost every ethical theory is some version of the principle of impartiality. According to this principle, our moral and morally relevant actions ought to result from objective and neutral deliberation, with all our personal interests, likes and dislikes left out. The intimate bond between the „impartial“ and the »moral« is not just an idea firmly rooted in our everyday moral discourse, but also a central tenet of many ethical theories, like the various „ideal“ or „impartial“ spectator theories advocated by, most notably, David Hume and Adam Smith. This principle figures prominently in the classic versions of utilitarian ethics, especially that advocated by John Stuart Mill [12], which will be central to the present consideration, as well as in its contemporary versions, such as that advanced by Peter Singer [17].

The basic idea of utilitarian ethics is that our actions must, if they are to count as morally right, bring about the greatest overall utility or happiness or the best consequences possible. Furthermore, when the utilitarian agent deliberates as to which action will yield the greatest utility or happiness or the best consequences, the impact of this action upon his own welfare must be regarded as irrelevant, or, at most, as equally important as its impact on all the other parties concerned. What *is* relevant from the utilitarian perspective is the maximization of good consequences or the amount of

* The first draft of this article was presented in December 2002 at the Konrad Lorenz Institute for Evolution and Cognition Research (Altenberg, Austria), as part of my research project on the evolutionary perspective on human nature and ethics. Its second draft was presented at a conference entitled „Freedom and Equality in Contemporary Philosophy“, organized by the Society for the Advancement of Philosophy in Samobor (Croatia) in June 2005. I thank both audiences for their constructive discussions.

happiness in general or on balance, *not* the maximization of one's own happiness. In the second chapter of his *Utilitarianism*, John Stuart Mill ([12], 64) says that „the happiness which forms the utilitarian standard of what is right in conduct, is not the agent's own happiness, but that of all concerned. As between his own happiness and that of others, utilitarianism requires him to be as strictly impartial as a disinterested and benevolent spectator“.

A vivid illustration of the possible impact of the utilitarian principle of impartiality upon moral practice was offered by William Godwin, the founder of philosophical anarchism and an early utilitarian. In his *Enquiry Concerning Political Justice* [7] Godwin asks that we imagine a house on fire, with Archbishop Fenelon and an ordinary chambermaid inside it. If we were able to save just one person, Godwin maintains, that person must be the archbishop, as his beneficial contribution to society at large is much greater than the chambermaid's. This decision is justified by virtue of its good consequences, and should be pursued regardless of any personal or partial relation we may have towards the chambermaid. In Godwin's words: „Supposing the chambermaid had been my wife, my mother or my benefactor. That would not alter the truth of the proposition. The life of Fenelon would still be more valuable than that of the chambermaid; and justice – pure, unadulterated justice – would still have preferred that which was most valuable. Justice would have taught me to save the life of Fenelon at the expense of the other. What magic is there, in the pronoun »my« to overturn the decisions of everlasting truth? My wife or my mother may be a fool or a prostitute, malicious, lying or dishonest. If they be, what consequence is it that they are mine?“ ([7], 41 – 42).

The principle of impartiality argued for by Mill and illustrated by Godwin is also recognizable in contemporary utilitarian approaches to ethics, like the one advocated by Peter Singer. At the center of Singer's „preference utilitarianism« stands one particular version of the principle of impartiality called „the principle of equal consideration of interests““. This principle, according to Singer, „acts like a pair of scales, weighing interests impartially. True scales favour the side where the interest is stronger or where several interests combine to outweigh a smaller number of similar interests; but they take no account of whose interests they are weighing“ ([17], 22). His well-known illustration of this is the obligation of donating money to overseas aid agencies. For example, the money we are about to spend on cars for our children, so that they can travel more easily to college, could, if donated to a humanitarian agency, prevent the death by hunger or infection of several children in some poor country. The interests of our children in having new cars definitely weigh less than the interests of starving or dying children. At this point, the utilitarian standard of morality, reinforced by the principle of impartiality, clearly indicates what our moral duties are. On the other hand, if we, following our affections and partiality, give preference to the persons we consider nearest and dearest, a determined utilitarian could accuse us of something quite close to murder; for, in this instance, we were able to save several lives, or at least one life, without sacrificing anything of comparable moral significance, yet we have failed to do so.

2. Evolutionary accounts of human partiality. Contrary to utilitarian requirements concerning general utility and impartiality, it is a well-known fact that human beings' moral and morally relevant decisions are usually highly partial: towards themselves, their

children and family members, their friends and neighbors, or their fellow-citizens, even towards members of their own race and sex. (Singer would also add members of their own species to this list.) In other words, people are not only highly reluctant to provide help to distant individuals whom they have never met nor will ever meet, but are also highly inclined to share their resources with a comparatively narrow circle of individuals with whom they are in some special intimate relation.

Human partiality is not just a permanent problem for the *normative* projects of almost every ethical theory; for a long time, it has also posed a problem for the *explanatory* projects of contemporary, evolution-oriented sciences regarding human behavior. In a nutshell, the problem was the following: how is it possible for natural selection, as the central mechanism of Darwinian evolution, to implant in humans *any* partial or altruistic behavioral tendencies towards anyone, if such behavior is diametrically opposed to the basic Darwinian principle of the survival of the fittest? This „paradox of altruism“ is captured well by Neven Sesardić ([15], 128): „[O]n one hand, it seems that the existence of human altruism is an undeniable psychological fact, but on the other hand it seems, on evolutionary grounds, that altruism cannot exist, because species with this trait are expected to have gone extinct through the process of natural selection. (Selfishness increases biological fitness, and only the fittest survive.)“¹

In spite of this *prima facie* conflict between altruistic behavior and the central tenets of Darwinian evolution, during the past three to four decades evolutionary approaches to human behavior – particularly sociobiology and evolutionary psychology – have managed to produce three currently quite influential explanations of human altruism.

The first explanation is the „kin selection“ or „inclusive fitness“ theory, originally proposed by William Hamilton [8]. According to this theory, human altruistic tendencies have evolved primarily in small groups of genetically related individuals. The basis of this theory is the idea of genes – »selfish genes«, to use the term coined by Richard Dawkins [6] – as crucial actors in the evolutionary process, whose primary »objective« is to create as many of their identical copies as possible in subsequent generations. From the perspective of genes which supposedly influence individual altruistic or non-altruistic behavior, it pays to be altruistic towards one's relatives because they carry a significant proportion of identical genes. By altruistically assisting the survival and reproduction of our relatives, we (i.e. our genes) are also assisting the survival and multiplication of copies of our genes. In this way, altruism evolved to become one of the most ubiquitous types of human social behavior.

The second explanation is the „reciprocal altruism“ theory, originally proposed by Robert Trivers [20]. Trivers' model does not necessarily require genetic relatedness between the performer and recipient of an altruistic act, but rather is designed „to show how certain classes of behavior conveniently denoted as ‚altruistic‘ (or ‚reciprocally altruistic‘) can be selected for even when the recipient is so distantly related to the organism performing the altruistic act that kin selection can be ruled out“ ([20], 189). According to this theory, altruistic tendencies evolve because it is advantageous to be altruistic towards individuals who are themselves altruistic. In this way, a cooperative network is created in which altruistic individuals, by virtue of their continuous mutual

¹ Sesardić [15] also offers an informative and innovative »state of the art« survey of theoretical approaches to human altruism and evolution during the past three to four decades.

aid, fare much better than non-cooperative individuals. Trivers' theory applies primarily (although not exclusively) to individuals with a longer life-span and more developed intelligence and memory. As humans definitely belong to this class, Trivers maintains that „given the universal and nearly daily practice of reciprocal altruism among humans today, it is reasonable to assume that it has been an important factor in recent human evolution and that the underlying emotional dispositions affecting altruistic behavior have important genetic components“ ([20], 212).²

The third explanation is the „group selection“ theory, originally proposed by Charles Darwin [5]. According to this theory, human altruistic tendencies have evolved by virtue of natural selection acting upon human groups or, as Darwin says, „tribes“. Although altruistic behavior may occasionally be disadvantageous for this or that particular individual, Darwin's point is that „a tribe including many members who [...] were always ready to aid one another, and to sacrifice themselves for the common good, would be victorious over most other tribes; and this would be natural selection“ ([5], 157 – 158). Darwin's idea of group-selected altruism was neglected in the human behavioral sciences for a long time, but was recently revived and improved – both conceptually and empirically – by Elliott Sober and David Sloan Wilson [19]. They convincingly argue that group selection is a plausible evolutionary force, and that much of human altruism – psychological altruism in particular – can be accounted for by means of a selection which acts not only on genes and individuals, but also on groups.

The theories of kin selection, reciprocal altruism and group selection are three more or less mutually consistent attempts to explain the commonsense fact that humans are especially altruistic or partial towards three classes of individuals: their relatives, individuals from whom they expect some kind of reciprocation, and members of their own group. Viewed from the opposite angle, the same theories explain why humans *do not* display altruism or partiality towards individuals who do not belong to any of these three classes. The point is that, in the course of evolution, investing one's resources in non-relatives, non-reciprocating individuals and group outsiders brought no bonus in terms of individual (inclusive) fitness and, as a result, humans remained psychologically ill-adapted (indisposed) to any such type of behavior.

If we accept these three explanations of human altruism or partiality as true, or at least as plausible, does this have any implications for the utilitarian principle of impartiality? In my view, some rather negative implications do indeed exist, and can be formulated by means of the two objections discussed below.

3. Objection from human evolutionary design. The first objection starts out from three standard points in Darwinian accounts of human nature: (1) human psychological mechanisms are shaped by natural selection and other causes of evolutionary change; (2) they are „designed“ as solutions to specific fitness-maximizing problems imposed on humans by their past physical and social environment; (3) just like all other biological

² Contemporary evolutionary psychologists like Leda Cosmides and John Tooby push this idea of Trivers further, arguing that the human mind is actually a series of mutually isolated modules or „Darwinian algorithms“, each of them specialized to solve a different problem posed to our Pleistocene ancestors by their „environment of evolutionary adaptedness“ (EEA). On the basis of their experiments with the „Wason selection task“, they argue in particular that the human mind contains a special „cheater detection module“ created by natural selection to facilitate reciprocally altruistic interactions ([2]; [3]).

traits, they are functionally constrained and cannot adequately perform tasks which differ significantly from those for which they were „originally designed“.

The first two points, as we have seen, allow us to predict that most humans will be motivated to sacrifice their own benefit for the sake either of their close relatives, or of individuals whom they perceive as possible partners in a prospective *cost-benefit* interaction, or of individuals whom they regard as members of their own group. Altruism is, therefore, a possibility in both evolutionary and psychological terms [19]. The third point, however, suggests that human altruism will occur in a restricted form and in specific circumstances only. As Elliott Sober and David Sloan Wilson emphasize, it is important to take into account the following »symmetric logic of selection arguments« concerning altruism and non-altruism: „... if selection promotes an altruistic concern for the welfare of one’s dear and near, it also may promote indifference or malevolence toward outsiders. Within-group selection is a competitive process, but so too is between-group selection; it promotes both within-group niceness and between-group nastiness. [...] If empathy elicits altruistic motives with respect to those whom we take to be similar, its absence means that we are less inclined to be altruistically motivated toward those whom we take to be different“ ([19], 326 – 327).

According to this »symmetric logic«, the class of individuals towards whom humans will not be altruistically motivated is much larger than the class of individuals towards whom they will. Selection for altruistic motives towards relatives, possibly reciprocating individuals, and one’s group insiders entails a lower degree, or even absence, of any similar motives towards non-relatives, individuals who are not likely to reciprocate, and group outsiders. Utilitarian impartiality, on the other hand, requires something completely different: equal benevolence to individuals from the latter three classes. Formulated in Godwinian terms: the utilitarian standard of rightness recognizes no »moral magic« in the fact that certain individuals are *my* relatives, *my* potential benefactors, or belong to *my* group.

In principle, a Darwinian might agree with this „no moral magic“ proviso, but he would probably then question the *raison d’être* of any moral standard which contradicts our best knowledge of human nature. In the case of utilitarian impartiality, this objection might assume the following form: since human psychology, due to its evolutionary „design“, is ill-adapted for producing behavior that impartially maximizes happiness, it is useless to proclaim such behavior to be morally obligatory. In its extreme version, this objection states that insisting on such behavior comes all too close to a violation of the „ought implies can“ principle. In its moderate version, it states that impartiality can be required in restricted form only, i.e. only in those social and psychological circumstances to which humans are relatively well-adapted.³

Any utilitarian would probably tend to dismiss the first (extreme) version of this

³ The same point can be formulated by analogy with human perceptual and reasoning capacities. As Gerhard Vollmer ([21], 88) suggests, „our sense organs, perceptual powers, structures of experience, ordinary language, and elementary inferential habits, are well adapted to this mesocosm [world of medium dimensions] and are *adequate* for mesocosmic needs“. However, these same organs and capacities do not function at the level of the microcosm or macrocosm (a world of very small or very large dimensions, respectively), as this is not the world in which they originated in the first place. In the same vein, it can be argued that human psychological capacities and motivational mechanisms cannot function well in social circumstances which differ significantly from those in which they first originated.

objection as too speculative or inconclusive. However, a realistic utilitarian would probably concede some weight to the second (moderate) version of this objection and attempt to devise a defense strategy. Mill, for example, suggested that the maximization of general happiness can and should be served first and foremost by promoting happiness at a more local level: „The multiplication of happiness is, according to the utilitarian ethics, the object of virtue: the occasions on which any person (except one in a thousand) has it in his power to do this on an extended scale, in other words, to be a public benefactor, are but exceptional; and on these occasions alone is he called on to consider public utility; in every other case, private utility, the interests or happiness of some few persons, is all he has to attend to“ ([12], 66).

However, in spite of its *prima facie* plausibility, this answer may perhaps only serve as a justification for the ethical views concerning impartiality held by the 19th century philosopher John Stuart Mill, but not for similar views held by his modern successors. It is more than obvious that Mill's world differs significantly from the modern world with respect to „occasions on which any person has it in his power to be a public benefactor“. These occasions are no longer „exceptional“, as Mill thought they were in his day; rather, due to possibilities offered by modern technology, they are almost omnipresent (at least for residents of Western societies). This creates a counter-intuitive situation which Michael Ruse ([14], 241) depicts in the following manner: „So long as one is dealing with a relatively tight little world, you can argue for treating most people more or less as equally deserving beings. [...] With a small, intimate group, all will benefit including yourself. But in the modern situation, to put utilitarianism into action simply leads to a never ending drain on your own resources, as you give everything away. And, somehow, there just does not seem to be the obligation to do this.“

As evolutionary psychologists like Leda Cosmides and John Tooby [33] argue, there is a significant discrepancy between human cultural and biological evolution. Cultural evolution, on the one hand, has done an amazing job: most humans no longer live as hunter-gatherers, in small kin-based groups, using stone-age technology, but rather in large cities, using computers and traveling in fast cars and airplanes. Biological evolution, on the other, has been less generous to humans: both physically and psychologically they are just the same as they were in their Pleistocene past. So even if human opportunities for public beneficence (primarily due to the technological possibilities they have at hand) are no longer »exceptional«, human motivation for this same beneficence (primarily due to constraints on their naturally selected psychology) remains „exceptional“ to a large extent.

A utilitarian might try to counter this objection by claiming that motivation by impartial considerations is not necessary for an action to qualify as moral; instead, all that is required is the *conformity* of this action with the standards of morality. This type of defense was offered by Mill, who argued that „the motive has nothing to do with the morality of the action“, so that, for example, „he who saves a fellow creature from drowning does what is morally right, whether his motive be duty, or the hope of being paid for his trouble“ ([12], 65). Mill's answer resembles the well-known »occasionalist« solution to the mind-body problem, according to which mind and body do not causally interact, but only appear to be causally connected because God constantly keeps them in mutual harmony. In a similar vein, there is, for Mill, no necessary connection between the

moral worth of an action and the human motives which bring it about, so long as its consequences are in line with utilitarian standards.

However, this »occasionalist« solution encounters some difficulties. For one thing, we usually expect moral actions to be motivated by genuine moral reasons. Otherwise, actions motivated by less than moral reasons and actions motivated by moral reasons proper would, if they accidentally led to the same final results, have the same worth. (For instance, I may give my entire property to a charity fund, motivated solely by my desire to leave a good impression on a jury which is about to convict me for stock-market fraud.) Mill's »occasionalist« solution thus runs counter to our common sense moral intuitions. Moreover, unlike the occasionalist solution to the mind-body problem, in which God synchronizes the independent spheres of the mental and the physical, the utilitarian ethical framework recognizes no similar instance that would warrant a presupposed ultimate synchronicity between the »motive of the action« and »morality of the action«. Finally, Mill himself points out that »no system of ethics requires that the sole motive of all we do shall be a feeling of duty«, adding that »ninety-nine hundredths of all our actions are done from other motives« ([12], 65). This however, in conjunction with the previous considerations, suggests that the principle of impartiality does not actually have any significant motivating power of its own, and may thus be considered irrelevant or redundant.

The »objection from human evolutionary design« amounts to claiming that humans generally lack the necessary motive to extend the boundaries of their natural benevolence. Unselfish motives available to humans will usually remain confined to particular classes of individuals, which are, however, too narrow from the perspective of utilitarian ethical standards. To be sure, this objection presupposes a Humean interpretation of moral motivation, whereby »reason is, and ought only to be the slave of the passions«, while »our sense of duty always follows the common and natural course of our passions« ([9], 462). A utilitarian may disagree with this emotivist conception, but then he is left with two independent burdens of proof: to prove both that a Humean view of moral motivation is wide of the mark, and that the principle of impartiality can play a significant role in actual moral practice.

4. Objection from natural human desires. The second objection relies on a common criticism of utilitarianism, one which suggests that the consistent maximization of general happiness would have such a detrimental effect on so many individual human qualities (both moral and non-moral) that its observance cannot be considered an overriding moral obligation. After a short illustration of two well-known formulations of this criticism by Bernard Williams and Susan Wolf, their possible link with Darwinian ideas will be addressed.

According to Bernard Williams ([23], 115), the utilitarian moral agent is no more than »the agent of the satisfaction system who happens to be at a particular point at a particular time.« His exclusive role is to serve as »a channel between the input of everyone's projects, including his own, and an output of optimific decision« (i.e. a decision that yields the best consequences in the given circumstances). Williams ([23], 116 – 117) maintains, however, that it is absurd to conceive human agents solely as members of one extensive system of the distribution and redistribution of general happiness, for in this

way one neglects the important fact that humans are separate individuals with their own personal integrity, as well as the fact that personal commitments, projects and goals play an indispensable role in their lives. Since utilitarianism thus threatens to alienate the very humanity and integrity of the moral agent, its principle of the impartial maximization of general happiness cannot be a moral principle or reason overriding all other moral or even extra-moral, reasons.

An objection similar to that of Williams is offered by Susan Wolf [24]. Wolf argues that the impartial maximization of general happiness as demanded by utilitarian ethical standards would amount to the sort of behavior characteristic of „moral saints“. The term »moral saint« bears no theological connotations here; rather, it describes a person for whom *moral* ideals, values and perfection always take precedence over any other human ideals, values and perfection. However, as Wolf claims ([24], 377), „moral perfection, in the sense of moral saintliness, does not constitute a model of personal well-being toward which it would be particularly rational or good or desirable for a human being to strive“. The problem with the cultivation of moral perfection is that it entails neglecting or totally sacrificing many other non-moral characteristics or activities which are usually considered to be irreplaceable parts of a well-rounded and fulfilled human life. What comes into play here again are quite personal projects and goals, such as advancing one's own talents and skills (intellectual, cultural, artistic, athletic, social and so on). If we devote our lives to the utilitarian furthering of general happiness, believes Wolf, we will necessarily have to sacrifice all these personal considerations which constitute the excellence of a person's life. If a moral theory ultimately entails sacrificing personal development, personal relations and a generally fulfilled life, then it may reasonably be opined that there is something wrong with such a theory.⁴

In connection with Darwinian accounts of human nature, these objections by Williams and Wolf seem particularly threatening to the viability of the utilitarian principle of impartiality. Namely, what Williams understands as „personal projects“, and Wolf as the ideal of the „fulfilled life“, corresponds significantly to a comparatively wide set of universal human inclinations that have been shaped over the course of biological evolution. The ethical bearing of those inclinations is particularly emphasized by Larry Arnhart (1998). Arnhart maintains that there are at least twenty *natural desires* present in all human societies throughout history, which constitute an inseparable part of an (ethically conceived) good life: a complete life, parental care, sexual identity, sexual mating, familial bonding, friendship, social ranking, justice as reciprocity, political rule, war, health, beauty, wealth, speech, practical habituation, practical reasoning, practical arts, aesthetic pleasure, religious understanding, and intellectual understanding ([1], 29).⁵

⁴ Moreover, as Wolf ([24], 381 – 382) remarks: „A world in which everyone, or even a large number of people, achieved moral sainthood – even a world in which they *strove* to achieve it – would probably contain less happiness than a world in which people realized a diversity of ideals involving a variety of personal and perfectionist values.“

⁵ A detailed exposition of Arnhart's [1] ethical views would require too much space here. It should be mentioned, however, that he attempts to combine the insights of contemporary Darwinism with certain classical views of human nature and human morality. He particularly defends the Aristotelian, and to some extent Humean, idea that human politics and morality are rooted in human biology, as well as the idea that there is no absolute division between natural facts and moral values. Although he does not mention Williams' and Wolf's points, many similarities do exist, especially concerning their conceptions of what constitutes a well-rounded and fulfilling human life.

It is not difficult to comprehend why many of these desires constitute inseparable parts of a distinctively human life. Parental care and sexual mating are *per definitionem* factors that maximize one's biological fitness; familial bonding, friendship and a preference for justice as reciprocity are socially important behavioral strategies, which have played a crucial role in human evolutionary history; and a preference for health and welfare is a practically essential condition for individual survival in the Darwinian »struggle for life«. It is important to bear in mind that, notwithstanding the fact of cultural evolution, these naturally evolved desires and preferences still represent central motivational mechanisms within the human behavioral and psychological repertoire. In other words, humans who lived as hunter-gatherers during the Pleistocene era and humans living today still have the same set of natural desires and the incentives to satisfy them.

The key words here are »naturally« and „personal“ and they indicate what seems to be at odds with the utilitarian ideal of impartiality. For it seems that the utilitarian principle of impartiality requires that an agent not only grant his own personal projects the same weight as those of any other agent, but also sacrifice many of his most personal lifelong projects. Mill, it should be recalled, explicitly states as follows: „As between his own happiness and that of others, utilitarianism requires him to be as strictly impartial as a disinterested and benevolent spectator“ ([12], 64). Even setting aside all Darwinian considerations, such a demand seems unjustified, as it tends to bring about consequences which are ethically absurd and psychologically unfavorable to the agent in question. However, this same objection derives additional force from a Darwinian theoretical framework, which demonstrates that human »personal projects« and visions of „a good life“ (which are usually supposed to be central to any ethical enterprise) actually consist of a series of natural, biologically rooted desires. It seems, therefore, that a serious attempt at leading a utilitarian moral life would not only require sacrificing some very important personal projects and ideals, but would also have a highly detrimental effect on the personal identity and character of the agent in question.

An apt illustration of this last point can perhaps be found in the life of John Stuart Mill himself. In his *Autobiography* [13] Mill describes his intellectual development, which was highly shaped, under the influence of his father, James Mill, and Jeremy Bentham, by utilitarian moral ideals and the desire to improve one's own analytic habits. However, as Mill points out, such an intellectual and emotional development resulted in a serious crisis in his mental history. As Mill later realized, analytic habits are „a perpetual worm at the root both of the passions and of the virtues; and above all, fearfully undermine all desires, and all pleasures...“ ([13], 97). He also notes that he was left „without any real desire« for the ends which he had been »so carefully fitted out to work for: no delight in virtue, or the general good, but also just as little in anything else“. The „fountains of vanity and ambition« had dried up in him, just as had »those of benevolence“ ([13], 98). Fortunately for Mill, and for the subsequent development of utilitarianism, his condition improved thanks to his pursuit of some rather personal projects and ideals (such as reading poetry). Although this story cannot be regarded as a knock-down argument against utilitarian claims, it at least suggests that the utilitarian principle of impartiality should not be taken at face value, especially when considered in conjunction with our best theories about human nature.

5. Concluding remarks. Why should utilitarianism be particularly vulnerable to any of the foregoing Darwinian objections? Are many other ethical theories not equally or even more susceptible to similar criticism? Theistic or Kantian ethics may likewise be considered irreconcilable with an evolutionary outlook; the former due to its assumption that the moral domain is somehow contingent upon God's will, the latter due to its assumption that the moral domain is not characterized by the causality of nature, but rather by a separate causality of freedom. To be sure, these ethical theories are also at odds with contemporary Darwinism, which allows for only one, natural type of causation, while the idea of any supernatural source of the human condition is practically dismissed from its theoretical horizon by default.

The relation between Darwinism and utilitarianism, however, is somewhat different, as they are both – unlike theism and Kantianism – committed to empiricism and naturalism. As Bernard Williams ([22], 83 – 85) appropriately remarks, (1) utilitarianism is „non-transcendental, and makes no appeal outside human life, in particular not to religious considerations“; (2) the basic good of utilitarianism (happiness) „seems minimally problematical: however much people differ, surely they at least all want to be happy“; (3) for utilitarians, „moral issues can, in principle, be determined by empirical calculation of consequences“; and (4) „utilitarianism provides a common currency of moral thought: the different concerns of different parties, and the different sorts of claims acting on one party, can all be cashed (in principle) in terms of happiness“.

Unlike some other influential ethical theories, it seems that utilitarianism, in its *descriptive* aspects, has almost borrowed its building blocks from the Darwinian theoretical inventory: in particular its non-transcendentalism, empiricism, and recognition of the human desire for happiness (especially if conceived in Benthamite terms of „seeking pleasure“ and „avoiding pain“). In its *normative* aspects, however, utilitarianism takes some additional steps that are not so easily accommodated into a general Darwinian framework. Why is this?

The key to this question probably lies in the utilitarian understanding of the idea of „human nature“. Mill, in line with his „associationism“ in psychology and the „moral sciences“ already proposed in his *System of Logic* [10], believes that human nature and the human „moral faculty“ are a „natural outgrowth“ from, but not a part of, our nature, which is malleable and perfectible to a large extent or, in his words, „susceptible of being brought by cultivation to a high degree of development“ ([12], 75). According to Mill, human nature is „susceptible, by a sufficient use of the external sanctions and of the force of early impressions, of being cultivated in almost any direction: so that there is hardly anything so absurd or so mischievous that it may not, by means of these influences, be made to act on the human mind with all the authority of conscience. To doubt that the same potency might be given by the same means to the principle of utility, even if it had no foundation in human nature, would be flying in the face of all experience ([12], 77).

Writing in *On Liberty*, Mill maintains ([11], 123) that human nature „is not a machine built after a model, and set to do exactly the work prescribed to it, but a tree, which requires to grow and develop itself on all sides, according to the tendency of the inward forces which make it a living thing“. Mill's optimistic and progressive picture of human nature was probably conceived as some sort of necessary psychological prerequisite for his strong utilitarian emphasis on the principle of impartiality. In this way, all that

remained to be done was to provide its *ethical* justification within the general utilitarian system.

However, Darwinian approaches such as sociobiology and evolutionary psychology suggest that such a psychological prerequisite for impartiality should not be accepted at face value. Theories like reciprocal altruism, kin selection and group selection predict that genuine impartiality should not be expected within the boundaries of human life, since the human capacity for benevolence on a socially extended scale is significantly constrained by human evolutionary design.⁶ As a historical point, moreover, it is interesting to note that even Darwin, who was Mill's contemporary, thought that the idea of the essential malleability of human moral behavior was not easily reconcilable with his own theory of human origins: „Mr. J. S. Mill speaks [...] of the social feelings as a ‚powerful natural sentiment‘, and as »the natural basis of sentiment for utilitarian morality.« Again he says, „Like the other acquired capacities above referred to, the moral faculty, if not a part of our nature, is a natural out-growth from it; capable, like them, in a certain small degree of springing up spontaneously.“ But in opposition to all this, he also remarks, »if, as in my own belief, the moral feelings are not innate, but acquired, they are not for that reason less natural.« It is with hesitation that I venture to differ at all from so profound a thinker, but it can hardly be disputed that the social feelings are instinctive or innate in the lower animals; and why should they not be so in man? [...] The ignoring of all transmitted mental qualities will, as it seems to me, be hereafter judged as a most serious blemish in the works of Mr. Mill ([5], 121).

Contemporary Darwinian perspective strongly suggests that utilitarian impartiality is equally and perhaps even more greatly endangered by two interconnected facts which have been considered in this article: (1) the fact that human biologically formed psychology and motivational mechanisms do not fit well with the principle requiring the indiscriminate promotion of general happiness; and (2) the fact that human beings are individuals with naturally evolved personal projects and ideals, and not merely the spare parts of some general utilitarian clockwork of happiness.

REFERENCES

- [1] ARNHART, L.: *Darwinian Natural Right: The Biological Ethics of Human Nature*. New York: State University of New York Press 1998.
- [2] COSMIDES, L.: The Logic of Social Exchange: Has Natural Selection Shaped How Humans Reason? Studies with the Wason Selection Task. In: *Cognition*, 31, 1989, p. 187 – 276.
- [3] COSMIDES, L. and TOOBY, J.: Cognitive Adaptations for Social Exchange. In: J. H. Barkow et al. (eds.): *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*.

⁶ It should be mentioned that some contemporary utilitarians are far more open to Darwinian considerations than Mill was. Peter Singer, for example, has written two relatively unbiased books about the ethical implications of contemporary Darwinism. Without downplaying any of the Darwinian considerations, he argues that the circle of human benevolence can, nevertheless, be expanded beyond biologically established human boundaries [16]. He also contends that the insights of evolutionary psychology concerning the human readiness for cooperation „should be seized upon as a starting point for the development of a field of social research that shows the way towards a more cooperative society“ ([17], 47). Singer's views are far too complex to be discussed in this limited space. Nevertheless, it could perhaps be argued that his interpretation of contemporary Darwinism remains too selective and, in particular, that he has lost sight of the »symmetric logic of selection arguments« mentioned in section 3 of this article.

Oxford: Oxford University Press 1992.

- [4] DARWIN, C.: *On the Origin of Species*. New York: Gramercy Books [1859], 1979.
- [5] DARWIN, C.: *The Descent of Man*. London, New York: Penguin [1871], 2004.
- [6] DAWKINS, R.: *The Selfish Gene*. Oxford, New York: Oxford University Press [1976], 1989.
- [7] GODWIN, W.: *Enquiry Concerning Political Justice and its Influence on General Virtue and Happiness*. New York: Knopf [1793], 1926.
- [8] HAMILTON, W. D.: The Evolution of Altruistic Behavior. In: T. H. Clutton-Brock and P. H. Harvey (eds.): *Readings in Sociobiology*. Reading and San Francisco: W. H. Freeman and Company [1963], 1978.
- [9] HUME, D.: *A Treatise of Human Nature*. London, New York: Penguin [1739 – 1740], 1985.
- [10] MILL, J. S.: *A System of Logic*. Honolulu: University Press of the Pacific [1843], 2002.
- [11] MILL, J. S.: *On Liberty*. London, New York: Penguin [1859], 1985.
- [12] MILL, J. S.: *Utilitarianism*. Oxford: Oxford University Press [1861], 2003.
- [13] MILL, J. S.: *Autobiography*. New York: Columbia University Press [1873], 1924.
- [14] RUSE, M.: *Taking Darwin Seriously*. New York: Prometheus Books [1986], 1998.
- [15] SESARDIĆ, N.: Recent Work on Human Altruism and Evolution. In: *Ethics*, 106, 1995, p. 128 – 157.
- [16] SINGER, P.: *The Expanding Circle: Ethics and Sociobiology*. Oxford, Melbourne: Oxford University Press 1981.
- [17] SINGER, P.: *Practical Ethics*. Cambridge: Cambridge University Press 1993.
- [18] SINGER, P.: *A Darwinian Left: Politics, Evolution and Cooperation*. New Haven and London: Yale University Press 2000.
- [19] SOBER, E. and WILSON, D. S.: *Unto Others: The Evolution and Psychology of Unselfish Behavior*. London: Harvard University Press 1998.
- [20] TRIVERS, R. L.: The Evolution of Reciprocal Altruism. In: T. H. Clutton-Brock and P. H. Harvey (eds.): *Readings in Sociobiology*. Reading and San Francisco: W. H. Freeman and Company [1971], 1978.
- [21] VOLLMER, G.: Mesocosm and Objective Knowledge – On Problems Solved by Evolutionary Epistemology. In: F. Wuketits (ed.): *Concepts and Approaches in Evolutionary Epistemology*. Dordrecht, Boston, Lancaster: D. Reidel 1984.
- [22] WILLIAMS, B.: *Morality: An Introduction to Ethics*. Cambridge: Cambridge University Press [1972], 2000.
- [23] WILLIAMS, B.: A Critique of Utilitarianism. In: J. J. C. Smart and B. Williams: *Utilitarianism: For & Against*. Cambridge: Cambridge University Press 1973.
- [24] WOLF, S.: Moral Saints. In: L. Pojman (ed.): *Ethical Theory: Classical and Contemporary Readings*. Wadsworth: Thomson Learning [1982], 2002.

Tomislav Bračanović
Departement of Philosophy
University of Zagreb – Croatian Studies
Ulica grada Vukovara 68
HR-10000 Zagreb
Chorvátsko (Croatia)
e-mail: tomislav.bracanovic@hrstud.hr
Tel. ++385 1 6117962