



OPEN ACCESS

EDITED BY

Sebastian Ille,
American University in Cairo, Egypt

REVIEWED BY

Elias L. Khalil,
Doha Institute for Graduate Studies, Qatar
Tomasz Kwarciński,
Uniwersytet Ekonomiczny w Krakowie, Poland

*CORRESPONDENCE

Geoffrey M. Hodgson
✉ g.hodgson2@lboro.ac.uk

RECEIVED 11 April 2025

ACCEPTED 04 August 2025

PUBLISHED 20 August 2025

CITATION

Hodgson GM (2025) Altruism and morality:
some problems for Max U.
Front. Behav. Econ. 4:1610022.
doi: 10.3389/frbhe.2025.1610022

COPYRIGHT

© 2025 Hodgson. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Altruism and morality: some problems for Max U

Geoffrey M. Hodgson*

Loughborough University London, London, United Kingdom

This essay considers some limitations of the assumption of utility maximization (Max U) as an explanation of human action. Leading promoters of the assumption explicitly avoid the question of motivation, to focus instead on manifest behavior. Max U is strictly unfalsifiable. With appropriate assumptions and contextual variables, it can be made to fit any real-world behavior. Max U is neither specifically human nor a product of evolution: it assumes a static set of preferences. Altruism, by definition, can be costly. If those costs result overall in a net disutility, even with a “warm glow” from helping others, then this is incompatible with the Max U assumption, where no voluntary behavior reduces utility. Relevant too are criticisms of “folk psychology,” where behavior is deemed to stem from stable preferences and beliefs. We also need to consider the nature of moral sentiments and how they evolved in cooperative human groups. Overall, an evolutionary perspective enables the development of a richer explanation of human behavior, including moral motivations, altruism and self-interested pursuits.

KEYWORDS

utility maximization, evolution, cooperation, altruism, moral sentiments

Introduction

Herbert Gintis made an enormous contribution to our understanding of human cooperation, including on how it has evolved and has been sustained. Since the 1990s, research into human cooperation has become a crowded field of theoretical and empirical research. Yet Herb (and his co-authors) used their immense intellectual skills to publish works that stand out and endure in this field.¹

There are many features of Gintis’s work that I admire. He recognized the importance of institutions in economics and acknowledged the power of evolutionary thinking. He argued that, under the right conditions, humans are a potentially cooperative, altruistic and moral creatures. But he and I differed on some theoretical issues, principally on the explanatory and theoretical value of the assumption that individuals are utility maximizers (Max U for short). He promoted this assumption. While distinguishing between self-regarding and other-regarding (social) preferences. Max U can encompass both, But I believe that is of little explanatory value. It focuses on behavior rather than underlying human nature. And crucially, it cannot adequately capture human moral motivation or altruism. At best, it offers a preliminary framework for more detailed, specific analysis.

1 See, for example, [Bowles and Gintis \(2005a,b\)](#); [Bowles and Gintis \(2011\)](#), [Boyd et al. \(2003\)](#), [Fehr and Gintis \(2007\)](#), [Gintis \(2000, 2003, 2004, 2009\)](#), [Gintis et al. \(2001, 2003, 2005\)](#), and [Gintis and Helbing \(2015\)](#). This article revises and updates some material from [Hodgson \(2013, 2014, 2015, 2021\)](#). The author is very grateful for the comments of three anonymous referees on a previous draft of this essay.

While I applaud Gintis's use of evolutionary theory, I argue here that it should be taken further, to try to explain the evolution of human moral capacities and how people develop, as their circumstances evolve. The Max U assumption can apply to any animal and even robots. While some economists may see such wide comprehensiveness as a virtue, the absence of uniquely human features makes it much more limited. The specificities of the human psyche and social organization must also be considered. In particular, we need evolutionary explanations of distinctively human dispositions. Following the evolutionary thinking of Veblen (1898, 1899) and others, this means explanations and discussions of motivation, rather than frameworks that merely fit or "express" behavior.

In this essay I outline some limitations of the Max U assumption. Eight sections follow. I point out that leading promoters of the Max U assumption explicitly avoid the question of motivation, to focus instead on manifest behavior. This leads to an explanation why Max U is strictly unfalsifiable. With appropriate assumptions and contextual variables, Max U can be made to fit any behavior. Rather than signaling explanatory power, this underlines its analytic weakness. Max U is neither specifically human nor a product of evolution: it assumes a static set of preferences. Subsequent sections discuss altruism and argue that, at least by its original definition, it is incompatible with Max U. Altruism, by definition, is costly. But Max U assumes that no behavior reduces utility. This leads to a critique of "folk psychology" where behavior is deemed to stem from stable preferences and beliefs. A subsequent section turns to morality, first by examining its nature or meaning. I report research since Darwin on how morality evolved. An evolutionary perspective enables the development of a richer view of human moral sentiments.

The explanatory emptiness of Max U

First, we must distinguish between payoff maximization and utility maximization (or Max U). A payoff is a reward in a game that has an explicit expected worth (such as a monetary reward) that is known to the analysts of the game and all its players.² Payoff maximization means the maximization of such explicit payoffs by players, given the information available to them, plus their assumption that other players are also payoff maximizers. Payoff rationality and self-interest were articles of faith among many mainstream economists from the 1950s to the 1990s. Behavioral economists have produced much evidence that challenges payoff rationality.³

Gintis is a co-author of a fascinating set of cross-cultural studies that show that players in ultimatum games rarely reach a Nash

payoff solution (Henrich et al., 2001; Henrich, 2004). One of two players in an ultimatum game is asked to divide an amount of money between herself and the other player. If the second player rejects the division, then both players get nothing; but if she accepts, then they each receive their allocated amounts. If the second player is a payoff maximizer, then she will accept the lowest possible positive allocation when it is offered: payoff maximizers always prefer something to nothing. This is a subgame perfect Nash equilibrium. But experiments often do not lead to this Nash payoff outcome: players do not always maximize payoffs in this way. These cross-cultural studies also showed that the actual pattern of play can vary significantly from one cultural setting to another.

Utility maximization is not necessarily payoff maximization, unless there is a monotonic relation between utilities and payoffs. This monotonic relationship is often assumed, but never in principle can it be demonstrated, because utility itself—even if it is inferred from behavior—is unobservable.

While undermining payoff maximization with experimental evidence, Gintis (2007, 2009) upheld consistency of behavior. Given behavioral consistency (or transitivity)—along with the other standard assumptions such as independence and continuity (Neumann and Morgenstern, 1944; Fishburn, 1970)—it is possible to construct an ordinal utility function where behavior is consistent with expected utility maximization. Gintis pointed out that a deviation from payoff maximization does not necessarily imply that people are behaving inconsistently or failing to maximize their utility.

But, instead of the "rational actor model," Gintis (2007, p. 2–3) preferred the following description: "I will generally refer to this as the beliefs, preferences, and constraints (BPC) model to avoid the often misleading connotations attached to the term 'rational.'" Consequently, unlike some other formulations of Max U, Gintis put beliefs first. He typically assumed core beliefs rather than explaining their origin or evolution.

Also, Gintis distinguished between "self-regarding" and "other-regarding" (or "social") preferences. With "self-regarding" preferences one takes account of one's own situation only. Gintis also distinguished between "self-regarding" and "self-interested" preferences. Hence for Gintis a charitable act is other-regarding rather than self-regarding, but it may be "self-interested" because of increased utility gained by the satisfying "warm glow" from helping or giving. Gintis (2006, p. 7) argued that "utility maximization should be a central tool in analyzing human behavior, even if humans are not self-regarding."

Gintis and his collaborators demoted the self-regarding preference function that dominated mainstream economic theory from the 1870s to the 1990s. They pointed to the experimental and other evidence that contradicts this narrow and selfish version of payoff maximization. They devised "other-regarding" preferences that are broadly consistent with the empirical evidence. But they rejected the claim that the empirical evidence falsifies Max U. Once we include the possibility of other-regarding preferences, and utility maximization that differs from payoff maximization, then Max U can be consistent with the empirical evidence.

In a joint paper, Gintis and Helbing (2015, p. 4–5) clarified their position: "The rational choice model expresses but does not explain individual preferences." They wrote: "the rationality assumption

² For simplicity I ignore games where possible payoffs are known to some players but not others. Including this possibility would not change the principal conclusions below.

³ See the extensive works of Nobel Laureate Daniel Kahneman and his colleagues (Kahneman, 1994, 2003a,b, 2011; Kahneman et al., 1982, 1986a,b). Bowles and Gintis (2011) provided an excellent overview of the behavioral evidence.

does not suggest that Alice is “trying” to maximize utility or anything else.” This depicts Max U as lacking in explanatory power concerning motivations or intentions. Other rational choice theorists have said something similar. For example, Posner (1980, p. 5) saw the “rationality of ‘economic man’” as “a matter of consequences, not states of mind.” Posner (p. 53 n.) continued: “in suggesting that people are economically rational, I am not making any statement about their conscious states. Rational behavior to an economist is a matter of consequences rather than intentions.”⁴ Posner, Gintis and Helbing may have disagreed on the degree to which preferences can be taken as “other-regarding,” but not on the non-explanatory, non-motivational, and non-intentional characterization of “rational” Max U.

Some version of Max U is very much a standard assumption within economics, but it has prominent dissenters. As Nobel Laureate Coase (1977, p. 488) remarked: “To say that people maximize utility tells us nothing about the purposes for which they engage in economic activity and leaves us without any insight into why people do what they do.” This suggests that Max U is inadequate as its explanation. Similarly, Nobel Laureate Sen (1977, p. 325) pointed to the circularity of explaining behavior “in terms of preferences, which are in turn defined only by behavior.” Sen (1987, p. 73) noted elsewhere that the description of choices in terms of utility “does not give any independent evidence on what the person is aiming to do or trying to achieve.” Both Coase and Sen agreed that Max U tells us little or nothing about motives or intentions.

The unfalsifiability of Max U

This leads to a related difficulty. When the young Samuelson (1937, p. 156) discussed utility maximization, he understood that “all types of observable behavior might conceivably result from such an assumption.” Because utility is unobservable, all kinds of behavior can be “explained” in terms of the idea, without fear of refutation. As Winter (1964, p. 228, 234) and Boland (1981) similarly remarked, no evidence can possibly refute the theory that agents are maximizing some hidden or unknown variable (such as utility).

Similarly to Gintis, while defending “self-interest, rightly understood” against its critics, Felin and Foss (2009, p. 662) said it is consistent with “cooperation, organization, community-building, trust, or for that matter, any other individual, relational, or organizational virtue.” Rather than selling the Max U assumption, this rather gives the game away. An assumption that is consistent with everything tells us very little, especially about specific motivational force.

Defenders of the Max U assumption might respond that inconsistent behavior would refute utility maximization. The problem here is one of identifying inconsistent behavior in empirical terms. Utility is unobservable. For example, if an experiment shows that option A with a value of \$4 is preferred to option B with a value of \$5 then we can simply assume that there are additional attributes of option A (for example, we may enjoy losing, or gain pleasure from seeing others win) that are

⁴ Posner (1980) also made the point that rationality (for him) does not necessarily imply conscious decision-making.

consistent with the view that it yields higher overall expected utility for the subject. Observable payoff non-maximization is “explained” by unobservable Max U.

On repeated visits to the same restaurant, we may prefer steak to fish 1 day, and fish to steak on another. Is this behavior inconsistent? Maybe. Maybe not. We may discover that the steak is not as good as expected. Or we may have seen an alarming television report about mad cow disease that causes us to switch to fish. The two choice occasions were different in terms of circumstances and knowledge. Hence, they do not necessarily imply inconsistency.

Given that we can never in principle demonstrate that some unobserved variable (like utility) is not being maximized, then the theory is invulnerable to any empirical attack. No amount of evidence can establish non-existence. Hence the standard core of expected utility theory is *unfalsifiable*.

But Max U is not a tautology in the logical sense because it is *conceivably false*. Logical tautologies—such as a triangle has three sides—are true by definition. If utility existed, then Max U might be true. Or it might be the case that individuals are not maximizing anything, even if their behavior seems consistent. These claims cannot be firmly established by empirical evidence. This does not necessarily mean that the utility maximization framework is useless or wrong. We do not have to uphold falsifiability as the mark of science—a criterion attributed to Karl Popper, whom in fact adopted a more nuanced position (Ackerman, 1976). Neither tautological nor non-falsifiable statements are necessarily meaningless or unscientific. They can be of scientific or heuristic value. They should be assessed in wider terms.⁵

A problem with utility maximization is that it can be adjusted to fit any behavior. Its explanatory success is an illusion. Close inspection of its proclaimed achievements reveal that the results depend on additional assumptions. For example, Nobel Laureate Becker (1976a, 1991, 1996) contended that standard rationality assumptions generate several testable predictions concerning human behavior. But all of Becker’s “predictions” depend on assumptions *additional* to his core assumption of utility maximization. Indeed, because it is difficult to conceive of evidence that falsifies these basic axioms, such models must depend on auxiliary assumptions to generate specific results (Pollak, 2003).

Orthodox economists may respond that Max U is essentially a simplifying assumption. It has proven its worth as an analytical tool and in making predictions about human behavior. But the claimed predictive success flows from the fact that Max U is unfalsifiable: preference functions can be contrived to fit any data. If we move from prediction to policy design, then there are severe limitations to an approach that cannot identify the specific underlying psychological and motivational forces involved. Without understanding the impulses behind behavior, the design of policies to guide behavior is in the dark.

⁵ Indeed, it is widely accepted in the philosophy of science—including by Popper—that some unfalsifiable propositions are necessary for science. These include the principle of determinacy (every event has a cause) and the assumption of the uniformity of nature. Without these prior assumptions, science is impossible.

Other arguments for retaining Max U do not involve claims of predictive or analytical success. It could be argued that Max U offers a universal analytical framework, or a starting point for more particular, context dependent analysis. These arguments hold some traction, and a fuller discussion is beyond the scope of this article. A key point here is that we would still have to move toward more specific, non-universal explanations of human phenomena such as altruism and moral motivation.

Max U is neither distinctively human nor a product of evolution

The notion of utility maximization is so capacious that it goes beyond humans. Experiments with rats and other animals (Kagel et al., 1981, 1995) led to suggestions that they have downward-sloping demand curves, just like humans. Becker (1991, p. 307) proposed that: “Economic analysis is a powerful tool not only in understanding human behavior but also in understanding the behavior of other species.” Similarly, Tullock (1994) has claimed that organisms—from bacteria to bears—can be treated as if they have the same general type of utility function that is attributed to humans in the microeconomics textbooks. Seemingly, we now have “evidence” of the rationality of everything in evolution from the amoeba onwards. But consequently, such assumptions are telling us very little about what is specific to human nature and human society.

For some economists, these extensions of the notion of utility maximization to the animal world are a triumphant vindication of the power and scope of orthodox economic analysis. But we should worry if the same utility-maximizing analysis is applied without modification to every organism. Such widely generalizable assumptions can tell us very little about what is specific to human nature and human society. Crucial and uniquely human aspects of our psychology are neglected. Consequently, there is no adequate and substantial theory of human agency or psychological motivation at the core of Max U. This weakness stems from its excessive scope. Purported explanations of everything can end up explaining little or nothing.⁶

Human societies are partly differentiated from other animals in terms of developed institutions and cultures. Max U bears no mark of any specifically human culture or institution. Human psychology is likewise neglected. Essentially, there is no adequate and substantial theory of *human* agency at the core of the standard theory. It tells us nothing of significance that is specific to the human psyche, human interaction, human nature, or human society. With respect to specifically human characteristics, it is causally vacuous.

Claims that there is an evolutionary basis for utility maximization (Robson, 2001a,b, 2002; Gintis, 2006, p. 17) do not pass muster. It is insufficient to show that the behavioral outcomes of evolution are consistent with some utility function. Ultimately this claim is trivially true, because one can always find a Max U

function that fits. One must show that Max U is a useful causal account of behavioral motivation. It is questionable that this has been achieved. A claim that X is a product of evolution where X can “explain” or “express” *any* behavior is of little use. Consequently, claims that Max U is a product of evolution have little significance.

Citing experimental evidence, some neuro-economists (Platt and Glimcher, 1999; Glimcher et al., 2005) have claimed that the utility function exists as a physiological reality inside the brain. This claim was scrutinized by Vromen (2010), who argued that at best the neurological evidence exhibits consistency with the predictions of expected utility theory. There is no evidence of the assessment of utility in the brain. Given the argument here that any observed outcomes can be made consistent with some utility function, the consistency claim is hardly powerful or surprising. But the existence claims are unsupported by the evidence.

It is important to understand the evolutionary origins of human dispositions. Max U serves as an ex-post rationalization or “expression” of behavior—rather than a grounded causal explanation. A utility function may have a limited purpose as a formalized preference ordering. Such formal constructions can be of some preliminary use. They can be shortcuts for modeling or explanatory purposes. But they do not enhance our understanding of human motivation, or of how human behavior is caused.

Max U's problem with altruism

Comte (1973, p. 566) coined the term *altruism*, meaning “living for others.” He saw it as the opposite of egoism, meaning the pursuit of one’s own desires. An altruist helps other people, with little reward or even at a personal cost. The philosopher Nagel (1970, p. 79) saw altruism as “a willingness to act in consideration of other persons, without the need of ulterior motives.” He explained further that altruism “will almost certainly include cases in which no obviously self-interested motive is present, and in which some inconvenience or at least no benefit to the agent is likely to result.” Based on a wide reading of the texts, William Scott Green (quoted in Wilson, 2015, p. 4) suggested a prevalent meaning of altruism as intentional “action ultimately for the welfare of others that entails at least the possibility of either no benefit or a loss to the actor.” Both Nagel and Green saw altruism as involving possible costs. Consequently, Nagel (ibid.) argued that a “defense of altruism in terms of self-interest is . . . unlikely to be successful.”

Even with other-regarding preferences, individuals are ultimately concerned simply with maximizing their own utility. Hence it is rather odd to claim simultaneously that (a) evolution has produced individuals who always maximize their own utility and (b) they are also capable of altruism. Altruism is typically defined as possibly costly for the individual concerned but beneficial for others. This sits uneasily with Max U. Consequently, within a utility-maximizing framework, the definition of altruism must be changed (e.g., Bowles and Gintis, 2011). Some Max U theorists working in an evolutionary framework awkwardly depict altruism as simultaneously involving a fitness cost and a utility gain for the agent, to preserve their enduring Max U claim.

Batson (2011) provided evidence of the strength of altruism in humans. In terms reminiscent of Smith (1759) notion of *sympathy*, he explained its persistence in terms of empathetic detection of

⁶ There are interesting and important phenomena, such as shame, self-deception and delusion, which may be problematic for a Max U approach (Khalil, 2016, 2017). Our main focus is on altruism and moral motivation. Hence, for present purposes, we may treat these issues as unresolved.

the needs of others. His experiments were carefully designed to distinguish between genuinely altruistic impulses and ulterior motives. Batson showed that when signs of strong empathy were present, self-serving explanations of altruism were not supported by evidence. Of course, proponents of Max U may see this as evidence for “other-regarding” preferences. That claim cannot be falsified. But the alternative is to attempt a deeper explanation by probing the Smithian interaction between sympathy and morality.

Zamagni (1995, p. xv, xvii) noted that some economists aim “to enlarge the scope of the rationality paradigm in economics,” to deal with the abundant empirical evidence of caring or “altruistic” behavior. According to these economists, “one need not step outside the homo economicus paradigm to accommodate altruism, provided self-interest is very broadly conceived and that one is prepared properly to expand the domain of the utility function of the rational decision maker.” Hence, to explain generosity, some economists assume that Max U gets additional “warm glow” utility by helping others (e.g., Becker, 1976b; Collard, 1978; Andreoni, 1990). The preference function is “other-regarding,” but Max U is still seeking to increase his own utility. Hence, as Zamagni (1995, p. xvii) argued, the problem with this modified Max U approach “is that it is not capable of capturing the true nature of altruism.” Even with “social” preferences, individual utility-maximization remains supreme. If an individual gains net utility by helping others, then he or she is still self-serving, rather than being altruistic in the established sense. Max U is still egoistically maximizing *his own* utility and pursuing *his own* self-interest. Modified Max U would not help others unless he gained utility from that behavior. Paradoxically this Max U, by helping others, is serving *his own selfishness more effectively*. This is not altruism, at least by its longstanding meaning.⁷

These economists violate the meaning of altruism to make it compatible with the Max U paradigm. Altruism now means doing something for others that increases the utility of the giver. It involves no net cost for the actor. Good Samaritans and other philanthropists are seen to maximize their own utility: at root they are all selfish, just like anyone else. In the eyes of economists, altruists remain egoists. Aldred (2019) argued, this is how the economics of Max U corrupts our understanding of human nature. Ultimately, whether we behave with overt generosity or greed, Max U brings us all down to the same selfish level.

Altruism beyond Max U

According to its original meaning, if people choose to act altruistically, then they are choosing to act with a possible personal (utility) cost to themselves. Even when the “warm glow” of utility gained (by seeing others benefit) is taken into account, altruism by definition involves a possible net cost. Strict utilitarians cannot accept this. Surely, rational agents must always be in search of some net benefit in terms of utility? To understand why choice can be costly, at least in terms of utility, we must abandon the view that we act in accord with well-defined preferences. Instead, human deliberation involves a tangled conflict of different beliefs, reasons, habits, instincts and emotions. The brain has no upper control

⁷ Note the argument by Wilson (2010) that “social preferences aren’t preferences.”

room that weighs the pros and cons of a considered action. Instead, it has a pre-frontal cortex that is pummel by feelings and impulses. We do choose. But, overwhelmed by internal urges and sensory data, the choosing mind is an emergent property of neural chaos in a complexly structured brain.

When a parent dives into a river to save his or her drowning child, the adult is driven by a mixture of urgent emotions and instincts. Reasons often come later: “I risked my life because saving my child was the right thing to do.” At least initially, emotions or moral sentiments can surpass risks and costs. At root, action derives neither from beliefs nor stable preferences. Instead, it is the outcome of tangled impulses and emotions, molded by specific contexts and triggered by events. Beliefs are involved, but these are empowered by activated emotions (Hodgson, 1988, ch. 5, 2006; 2010; Damasio, 1994). Reflection may come later, when moral commitments may then be nuanced or reinforced.

Psychological experiments show that conscious sensations are reported about half a second after neural events, and unconscious brain processes are discernible before any conscious decision to act. This suggests that our dispositions are activated before our actions are rationalized: we contrive reasons for actions already under way. This undermines the “folk psychology” that actions are primarily driven by reasons or beliefs. This “folk psychology” supports “mind first” explanations of human behavior that are unable to explain adequately such phenomena as sleep, memory, learning, mental illness, or the effects of chemicals or drugs on our perceptions or actions.⁸

Once we abandon belief-first and mind-first notions of motivation, feelings and emotions become more important, and the subjective costs of altruism can be understood. Something must give—either the established meaning of altruism or the assumption of Max U. Doing something for others, when the utility of the doer is reduced, is impossible within a Max U explanatory framework. Max U never willingly suffers a loss of utility, even when helping others.

We must look at thoughts and feelings, as well as consequences and behavior. The Good Samaritan was aware of the lethal ethnic divisions in the locality, and the evident risks of robbery on the road from Jerusalem to Jericho. It would have been wiser to hurry on to his destination. But, knowing these risks, he was overwhelmed by sympathy and looked after the roadside victim. The parent was acutely aware of the severe risks in diving into a freezing and swollen river to save a child. But parental and compassionate feelings won out. In both cases, substantial risks or costs were perceived but pushed aside. That is why their actions were altruistic.

The nature of morality

Smith (1759, p. 188) argued that sympathy and moral sentiments worked alongside self-interest, and it was vital to take all of these on board. For Smith, utility was “not the first or principal source of our approbation ... the sentiment of approbation always

⁸ On the experimental evidence see Libet (1985); Libet et al. (1983); Wegner (2002, 2003); Wegner and Wheatley (1999); Haynes and Rees (2005a,b); Haynes et al. (2007). On the limits of “belief first” or “mind first” explanations see Bunge (1980), Stich (1983), Churchland (1984, 1986), Damasio (1994); Rosenberg (1998).

involves in it a sense of propriety quite distinct from the perception of utility.” For Smith, moral or other virtue was irreducible to utility.

Drawing on what I believe is the majority view among moral philosophers, I maintain that *moral dispositions cannot be adequately summarized by a (Max U) preference function*. As noted above, preference functions can be contrived to fit any behavior. In this shallow sense of an empirical fit, a preference function can “express” moral motivation. But it adds nothing to our understanding of our evolved moral dispositions and the role of emotion in sustaining them. The language of “preferences” is inadequate to express the nature of morality.

Morality is complex and controversial. In Darwin’s (1871, vol. 1, p. 87–89) account, morality results from a combination of emotional impulses and thoughtful deliberation. He argued that although primitive moral feelings have evolved for millions of years among “the progenitors of man” (Darwin, 1871 p. 162), humans alone have a developed sense of morality:

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

For Darwin, morality emerged in humans upon a long-evolved foundation of instinct and impulse. Darwin also saw morality as a social phenomenon, involving social relations and shared values. Many controversies divide moral philosophers. The best we can do here is to select a few prominent descriptions of the nature of moral judgment. The moral philosopher Hare (1952) argued that morality was subject to reason, and one cannot hold contradictory ethical judgments. As Mackie (1977, p. 33) put it in his account, a moral judgment:

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

In his important philosophical treatise on the *Evolution of Morality*, Joyce (2006, p. 70) drew from the literature in ethics and argued that morality has most or all of the following characteristics:

- Moral judgments express attitudes (such as approval or contempt) and also express beliefs.
- The emotion of guilt is an important mechanism for regulating moral conduct.
- Moral judgments transcend the interests or ends of those concerned.
- Moral judgments imply notions of desert and justice.
- Moral judgments are inescapable.
- Moral judgments transcend human conventions.
- Moral judgments govern interpersonal relations and counter self-regarding individualism.

These characteristics do not establish a *valid* morality; they instead help us to identify what is a *moral judgment*, whether acceptable or otherwise. The argument here relies on descriptive

rather than normative ethics: there is no attempt here to identify the “right” morality, but instead to identify the basic nature of a moral claim. Most religions uphold moral claims, but that does not make them all right or just.

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

Moral judgments may be rationalized in various ways, but they are more than matters of belief. The emotional dimension of moral rules plays an important role in their evolution and their survival. Guilt [point (2)] is a particularly important emotion that sometimes emerges after breaches of moral rules, and it too plays a part in the evolutionary process.

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

As both Mackie and Joyce insisted, morality surpasses questions of preference. It is a matter of right or wrong, or of duty, of “doing the right thing,” irrespective of whether we like it or not. This is part of what makes us human: we can consider moral rules and understand that their observance is more than a matter of personal whim or satisfaction. This dimension is missing in much of economics. Moral values are subsumed under matters of utility or preference. This has important consequences for policy design (Hodgson, 2013).

Modern society establishes a difference between moral rules and other (normative) rules. Linguistic and traffic rules are matters of convention; they are non-universal. Murder is also punishable, but it is much more than a breach of convention. Threat of punishment or respect for the law are each insufficient to explain the relatively low frequency of murder and other immoral crimes. Most of us abstain from murder not simply because the probability of severe punishment outweighs any expected benefit. Most of us refrain from murder because we believe that it is *morally wrong*; we would desist even if we lived in a country where murder went unpunished. Tom Tyler and his colleagues have gathered considerable evidence to show that people often obey the law because it is morally right, not generally because they calculate the costs and benefits of doing so (Tyler, 2006, 2017).

In summary, moral judgments involve expressions of attitudes, beliefs and emotions but are also subject to deliberation concerning matters of fairness or justice. In contrast to standard utilitarian approaches, a moral judgment is more than mere convention; it is inescapable and transcends individual preferences or interests (Smart and Williams, 1973). Of course, when faced with moral dilemmas, people do often weigh up one option against the other. But to describe all this as utility maximization or a matter of preference misses the point. If people are always acting in a way that ends up maximizing *their own* utility, then they cannot be seen as truly altruistic or wholeheartedly moral.

The evolution of morality

Most advocates of the Max U approach fail to explain how Max U evolved, and the few attempts to explain this are unconvincing because Max U applies to any behavior. Evolutionary explanations must attempt to explain specific phenomena, not unfalsifiable generalities that fit everything and explain very little. The literature on this starts with Darwin and has experienced a major revival of research and analysis since the 1990s. There are several important disputes and only a partial consensus on key issues. But the minimum claim is this: assumptions about human motivation should be consistent with a credible understanding of human evolution. We require plausible evolutionary explanations of the origin and persistence of morality.

How did human cooperation evolve? Darwin's answer is remarkably resilient, even in the light of modern research. Darwin was unaware of genetics, but in his *Descent of Man*, he considered how "sympathy, fidelity, and courage" would advantage one group against the other in their struggle for existence. Darwin (1871, vol. 1, p. 162) considered such traits, which had been originally "acquired by the progenitors of man." He noted the disposition to obey the commands and rules of those in authority rather than to always follow individually selfish motives:

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

Darwin (1871, vol. 1, p. 166) also argued that "although a high standard of morality gives but a slight or no advantage to each individual man and his children over the other men of the same tribe," the advancement and enforcement of moral rules would "give an immense advantage to one tribe over another." Tribes with members possessing a high degree "of patriotism, fidelity, obedience, courage, and sympathy" and "ready to give aid to each other and to sacrifice themselves for the common good, would be victorious over most other tribes; and this would be natural selection." Darwin proposed that groups containing individuals that followed social rules that served the interests of their group will have an advantage in the struggle for survival.

Some groups are more successful than others. For much of human existence, competition between groups has often involved violent combat. More cooperative groups, with cohering rules and norms, won out over other groups. But Darwin did not explain how altruistic or cooperative traits that favored group survival emerged within a group and spread within it. His account of group selection was highly perceptive but incomplete. We have now much more evidence and additional robust arguments to support this conjecture. Evidence also shows the endurance of violent conflict from prehistoric times. These enhanced the pressures of group selection (Bowles, 2009; Lawler, 2012; Nivette, 2011; Pinker, 2011; Turchin, 2016).

If some individuals in a group were disposed to cooperate, what would stop them being overwhelmed by free riders? Free riders might already exist or develop within the group, or they might be allowed in from outside. They would benefit from cooperation and altruism but fail to contribute. As free riders,

they would not bear the costs of cooperation or altruism, but they would enjoy the benefits. Within the group, free riders would prosper and eventually outnumber the cooperators. The outcome would be deleterious for the group, but it would not stop the triumph of the selfish. These arguments were framed in gene-centered terms in the 1960s and 1970s by George C. Williams and most famously by Richard Dawkins. Genes promoting free riding would be more successful than genes promoting cooperation or altruism. Darwin's claims about group selection seemed unsound (Williams, 1966; Dawkins, 1976).

Darwin's argument relied on a notion of group selection, which recently has been rehabilitated and refined (Sober and Wilson, 1998; Henrich, 2004; Turchin, 2016). We now have rigorous accounts of possible mechanisms of group selection, including the important distinction between genetic and cultural group selection. The relationships and interactions between genetic and cultural factors in the evolution of morality are under dispute, but perhaps the best guess (based in current evidence) is that the genetic foundations of altruistic and moral feelings are likely to have evolved first through mechanisms of kin altruism and then reciprocal altruism. Then altruistic, cooperative and moral feelings required the further emergence of a culture, so that they could spread further through the group and become reinforced by enduring cultural norms. In short, genetic mechanisms established critical masses of altruists in social groups, leading to the spread of cultural norms sustaining cooperation and to the development of systems of morality that further enhanced the fitness of groups (Hodgson, 2013). The foundations of morality evolved first on a kin-based genetic basis, then morality was formed and transmitted culturally within groups. Groups with stronger moral ties won out in competition with other groups.⁹

Conclusion

Economists claim that the Max U assumption has been successful in providing predictions of human behavior. But apparent success in prediction does not imply success in understanding underlying mechanisms or designing policies. Institutional and policy design require an adequate understanding of individual motivation. If we conceive individuals simply as pleasure machines, then we downplay their moral sentiments. We appeal to their pleasure, and not directly to their morality or sense of duty. With a grossly insufficient understanding of real motives, the derived policies are likely to be flawed.

We must look more closely at the details, including some recent versions of Max U, including those based on "other-regarding" preferences. Do these adequately accommodate the notion of morality? Utilitarianism is a version of consequentialism—this means that acts are appraised ethically in terms of their outcomes alone. Consequentialism is not confined to mainstream economics. But there are problems with its philosophical and ethical underpinnings. As Kant (2002) argued long ago, an ethical focus on consequences is not enough.

⁹ Several authors have argued that religion was important in framing and promoting moral sentiments in human groups (Norenzayan, 2013; Johnson, 2016).

Tolerably accurate imputation of motives is vital for social interaction and cooperation. Discernment of sufficient integrity and commitment is essential to engender trust. It is important to detect those who are unreliable or less sincere. Simply acting with apparent altruism or morality is not enough. If actions are insincere, then they are unreliable in other circumstances. Overt behavior is a poor guide. Accurate imputations of moral integrity (or otherwise) matter.

An understanding of these issues is vital for the design of policies to deal with tax compliance, corruption, public finance, health systems, care for those in need, climate change, and so on. If we assume that everyone is maximizing their own utility, even if they get a “warm glow” of extra utility from helping others, then we have a highly imperfect basis for policy design. We focus on consequences rather than the cultivation of inner motives. We neglect the massive evidence from psychology that humans are much more complicated and sophisticated than Max U.

Theorists like Herbert Gintis and others are to be commended for bringing issues like altruism and cooperation to the forefront. We need to go further, and attend to the problems with some of the underlying assumptions in standard economic analysis.

Author contributions

GH: Conceptualization, Methodology, Writing – original draft.

References

- Ackerman, R. (1976). *The Philosophy of Karl Popper*. Amherst, MA: University of Massachusetts Press.
- Aldred, J. (2019). *Licence to be Bad: How Economics Corrupted Us*. London: Allen Lane.
- Andreoni, J. (1990). Impure altruism and donations to public goods: a theory of warm-glow giving. *Econ. J.* 100, 464–477. doi: 10.2307/2234133
- Batson, C. D. (2011). *Altruism in Humans*. Oxford; New York, NY: Oxford University Press.
- Becker, G. S. (1976a). *The Economic Approach to Human Behavior*. Chicago, IL: University of Chicago Press.
- Becker, G. S. (1976b). Altruism, egoism, and genetic fitness: economics and sociobiology. *J. Econ. Literat.* 14, 817–826.
- Becker, G. S. (1991). *A Treatise on the Family, 2nd Edn.* Cambridge, MA: Harvard University Press. doi: 10.4159/9780674020665
- Becker, G. S. (1996). *Accounting for Tastes*. Cambridge, MA: Harvard University Press. doi: 10.4159/9780674020658
- Boland, L. A. (1981). On the futility of criticizing the neoclassical maximization hypothesis. *Am. Econ. Rev.* 71, 1031–1036.
- Bowles, S. (2009). Did warfare among ancestral hunter-gatherer groups affect the evolution of human social behaviors? *Science* 324, 1293–1298. doi: 10.1126/science.1168112
- Bowles, S., and Gintis, H. (2005a). *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life*. Cambridge, MA: MIT Press.
- Bowles, S., and Gintis, H. (2005b). Can self-interest explain cooperation? *Evol. Instit. Econ. Rev.* 2, 21–41. doi: 10.14441/eier.2.21
- Bowles, S., and Gintis, H. (2011). *A Cooperative Species: Human Reciprocity and its Evolution*. Princeton, NJ: Princeton University Press. doi: 10.23943/princeton/9780691151250.001.0001
- Boyd, R., Gintis, H., Bowles, S., and Richerson, P. J. (2003). Evolution of altruistic punishment. *Proc. Natl. Acad. Sci.* 100, 3531–3535. doi: 10.1073/pnas.0630443100
- Bunge, M. A. (1980). *The Mind-Body Problem: A Psychobiological Approach*. Oxford: Pergamon. doi: 10.1016/B978-0-08-024720-5.50006-2
- Churchland, P. M. (1984). *Matter and Consciousness*. Cambridge, MA: MIT Press.
- Churchland, P. S. (1986). *Neurophilosophy: Toward a Unified Science of the Mind-Brain*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/4952.001.0001
- Coase, R. H. (1977). “Economics and contiguous disciplines,” in *The Organization and Retrieval of Economic Knowledge*, ed. M. Perlman (Boulder, CO: Westview Press), 481–491. doi: 10.1007/978-1-349-03325-6_26
- Collard, D. (1978). *Altruism and Economy: A Study in Non-Selfish Economics*. Oxford: Martin Robertson.
- Comte, A. (1973). *System of Positive Polity, Volume 1: Containing the General View of Positivism and Introductory Principles*, translated by John Henry Bridges from the French edition of 1851. New York, NY: Burt Franklin.
- Damasio, A. R. (1994). *Descartes Error: Emotion, Reason, and the Human Brain*. New York, NY: Putnam.
- Darwin, C. R. (1871). *The Descent of Man, and Selection in Relation to Sex, 1st Edn., 2 Vols*. London: John Murray; New York, NY: Hill. doi: 10.1037/12294-000
- Dawkins, R. (1976). *The Selfish Gene, 1st Edn.* Oxford: Oxford University Press.
- Fehr, E., and Gintis, H. (2007). Human motivation and social cooperation: experimental and analytical foundations. *Annu. Rev. Sociol.* 33, 43–64. doi: 10.1146/annurev.soc.33.040406.131812
- Felin, T., and Foss, N. J. (2009). Social reality, the boundaries of self-fulfilling prophecy, and economics. *Organiz. Sci.* 20, 654–668. doi: 10.1287/orsc.1090.0431
- Fishburn, P. C. (1970). *Utility Theory for Decision Makers*. New York, NY: Wiley. doi: 10.21236/AD0708563
- Gintis, H. (2000). Strong reciprocity and human sociality. *J. Theoret. Biol.* 206, 169–179. doi: 10.1006/jtbi.2000.2111
- Gintis, H. (2003). The hitchhikers guide to altruism: genes, culture, and the internalization of norms. *J. Theoret. Biol.* 220, 407–418. doi: 10.1006/jtbi.2003.3104

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Gintis, H. (2004). The genetic side of gene-culture coevolution: internalization of norms and prosocial emotions. *J. Econ. Behav. Organiz.* 53, 57–67. doi: 10.1016/S0167-2681(03)00104-5
- Gintis, H. (2006). Behavioral ethics meets natural justice. *Polit. Philos. Econ.* 5, 5–32. doi: 10.1177/1470594X06060617
- Gintis, H. (2007). A framework for the integration of the behavioral sciences. *Behav. Brain Sci.* 30, 1–16. doi: 10.1017/S0140525X07000581
- Gintis, H. (2009). *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences*. Princeton, NJ: Princeton University Press.
- Gintis, H., Bowles, S., Boyd, R., and Fehr, E. (2003). Explaining altruistic behavior in humans. *Evol. Hum. Behav.* 24, 153–172. doi: 10.1016/S1090-5138(02)00157-5
- Gintis, H., Bowles, S., Boyd, R., and Fehr, E. (2005). *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/4771.001.0001
- Gintis, H., and Helbing, D. (2015). Homo socialis: an analytical core for sociological theory. *Rev. Behav. Econ.* 2, 1–59. doi: 10.1561/105.000000016
- Gintis, H., Smith, E. A., and Bowles, S. (2001). Costly signaling and cooperation. *J. Theoret. Biol.* 213, 103–119. doi: 10.1006/jtbi.2001.2406
- Glimcher, P. W., Dorris, M. C., and Bayer, H. M. (2005). Physiologic utility theory and the neuroeconomics of choice. *Games Econ. Behav.* 52, 213–256. doi: 10.1016/j.geb.2004.06.011
- Hare, R. M. (1952). *The Language of Morals*. Oxford: Oxford University Press.
- Haynes, J.-D., and Rees, G. (2005a). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nat. Neurosci.* 8, 686–691. doi: 10.1038/nn1445
- Haynes, J.-D., and Rees, G. (2005b). Predicting the stream of consciousness from activity in human visual cortex. *Curr. Biol.* 15, 1301–1307. doi: 10.1016/j.cub.2005.06.026
- Haynes, J.-D., Sakai, K., Rees, G., Gilbert, S., Frith, C., and Passingham, R. E. (2007). Reading hidden intentions in the human brain. *Curr. Biol.* 17, 323–328. doi: 10.1016/j.cub.2006.11.072
- Henrich, J. (2004). Cultural Group selection, coevolutionary processes and large-scale cooperation. *J. Econ. Behav. Organiz.* 53, 3–35. doi: 10.1016/S0167-2681(03)00094-5
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., et al. (2001). In search of homo economicus: behavioral experiments in 15 small-scale societies. *Am. Econ. Rev.* 91, 73–84. doi: 10.1257/aer.91.2.73
- Hodgson, G. M. (1988). *Economics and Institutions: A Manifesto for a Modern Institutional Economics*. Cambridge, MA; Philadelphia, PA: Polity Press and University of Pennsylvania Press.
- Hodgson, G. M. (2006). Instinct and habit before reason: comparing the views of John Dewey, Friedrich Hayek and Thorstein Veblen. *Adv. Austrian Econ.* 9, 109–143. doi: 10.1016/S1529-2134(06)09005-3
- Hodgson, G. M. (2010). Choice, habit and evolution. *J. Evol. Econ.* 20, 1–18. doi: 10.1007/s00191-009-0134-z
- Hodgson, G. M. (2013). *From Pleasure Machines to Moral Communities: An Evolutionary Economics Without Homo Economicus*. Chicago, IL: University of Chicago Press. doi: 10.7208/chicago/9780226922737.001.0001
- Hodgson, G. M. (2014). The evolution of morality and the end of economic man. *J. Evol. Econ.* 24, 83–106. doi: 10.1007/s00191-013-0306-8
- Hodgson, G. M. (2015). A trojan horse for sociology? Preferences versus evolution and morality. *Rev. Behav. Econ.* 2, 93–112. doi: 10.1561/105.000000021
- Hodgson, G. M. (2021). *Liberal Solidarity: The Political Economy of Social Democratic Liberalism*. Cheltenham; Northampton, MA: Edward Elgar. doi: 10.4337/9781800882171.00005
- Johnson, D. (2016). *God is Watching You: How the Fear of God Made Us Human*. Oxford; New York, NY: Oxford University Press.
- Joyce, R. (2006). *The Evolution of Morality*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/2880.001.0001
- Kagel, J. H., Battalio, R. C., and Green, L. (1995). *Economic Choice Theory: An Experimental Analysis of Animal Behaviour*. Cambridge, MA; New York, NY: Cambridge University Press. doi: 10.1017/CBO9780511664854
- Kagel, J. H., Battalio, R. C., Rachlin, H., and Green, L. (1981). Demand curves for animal consumers. *Quart. J. Econ.* 96, 1–16. doi: 10.2307/2936137
- Kahneman, D. (1994). New challenges to the rationality assumption. *J. Instit. Theoret. Econ.* 150, 18–36.
- Kahneman, D. (2003a). A psychological perspective on economics. *Am. Econ. Rev.* 93, 162–168. doi: 10.1257/000282803321946985
- Kahneman, D. (2003b). Maps of bounded rationality: psychology for behavioral economics. *Am. Econ. Rev.* 93, 1449–1475. doi: 10.1257/000282803322655392
- Kahneman, D. (2011). *Thinking, Fast and Slow*. London; New York, NY: Penguin.
- Kahneman, D., Knetsch, J. L., and Thaler, R. H. (1986a). Fairness as a constraint on profit seeking: entitlements in the market. *Am. Econ. Rev.* 76, 728–741.
- Kahneman, D., Knetsch, J. L., and Thaler, R. H. (1986b). Fairness and the assumptions of economics. *J. Business* 59, 285–300. doi: 10.1086/296367
- Kahneman, D., Slovic, P., and Tversky, A. (eds.) (1982). *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge; New York, NY: Cambridge University Press.
- Kant, I. (2002). *Groundwork for the Metaphysics of Morals, translated from the German edition of 1785*. New Haven, CT; London: Yale University Press.
- Khalil, E. L. (2016). Self-deception as a weightless mask. *Facta Univ. Ser.: Philos. Sociol. Psychol. Hist.* 15, 1–11.
- Khalil, E. L. (2017). Making sense of self-deception: distinguishing self-deception from delusion, moral licensing, cognitive dissonance and other self-distortions. *Philosophy* 92, 539–563. doi: 10.1017/S003181917100033X
- Lawler, A. (2012). Civilizations double-edged sword. *Science* 336, 832–833. doi: 10.1126/science.336.6083.832
- Libet, B. (1985). Unconscious Cerebral initiative and the role of conscious will in voluntary action. *Behav. Brain Sci.* 8, 529–566. doi: 10.1017/S0140525X00044903
- Libet, B., Gleason, C. A., Wright, E. W., and Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): the unconscious initiation of a freely voluntary act. *Brain* 106, 623–642. doi: 10.1093/brain/106.3.623
- Mackie, J. L. (1977). *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin.
- Nagel, T. (1970). *The Possibility of Altruism*. Oxford; Princeton, NJ: Clarendon Press and Princeton University Press.
- Neumann, J., and Morgenstern, O. (1944). *The Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.
- Nivette, A. E. (2011). Violence in non-state societies. *Br. J. Criminol.* 51, 578–598. doi: 10.1093/bjc/azr008
- Norenzayan, A. (2013). *Big Gods: How Religion Transformed Cooperation and Conflict*. Princeton, NJ: Princeton University Press. doi: 10.1515/9781400848324
- Pinker, S. (2011). *The Better Angels of Our Nature: Why Violence Has Declined*. New York, NY: Viking.
- Platt, M. L., and Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238. doi: 10.1038/22268
- Pollak, R. A. (2003). Gary Beckers contributions to family and household economics. *Rev. Econ. Household* 1, 111–141. doi: 10.1023/A:1021803615737
- Posner, R. A. (1980). A theory of primitive society, with special reference to law. *J. Law Econ.* 23, 1–53. doi: 10.1086/466951
- Robson, A. J. (2001a). The biological basis of human behavior. *J. Econ. Literat.* 39, 11–33. doi: 10.1257/jel.39.1.11
- Robson, A. J. (2001b). Why would nature give individuals utility functions? *J. Polit. Econ.* 109, 900–914. doi: 10.1086/322083
- Robson, A. J. (2002). Evolution and human nature. *J. Econ. Perspect.* 16, 89–106. doi: 10.1257/0895330027274
- Rosenberg, A. (1998). “Folk psychology,” in *Handbook of Economic Methodology*, eds. J. B. Davis, D. W. Hands, and U. Mäki (Cheltenham: Edward Elgar), 195–197.
- Samuelson, P. A. (1937). A note on the measurement of utility. *Rev. Econ. Stud.* 4, 155–161. doi: 10.2307/2967612
- Sen, A. K. (1977). Rational fools: a critique of the behavioral foundations of economic theory. *Philos. Public Affairs* 6, 317–344.
- Sen, A. K. (1987). “Rational behaviour,” in *The New Palgrave Dictionary of Economics*, Vol. 4, eds. J. Eatwell, M. Milgate, P. Newman (London: Macmillan), 68–76.
- Smart, J. J. C., and Williams, B. (1973). *Utilitarianism: For and Against*. Cambridge, MA: Cambridge University Press.
- Smith, A. (1759). *The Theory of Moral Sentiments; or, An Essay Towards an Analysis of the Principles by which Men Naturally Judge Concerning the Conduct and Character, First of their Neighbours, and Afterwards of Themselves*. London; Edinburgh: Millar, and Kincaid and Bell.
- Sober, E., and Wilson, D. S. (1998). *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press.
- Stich, S. P. (1983). *From Folk Psychology to Cognitive Science*. Cambridge, MA: MIT Press.
- Tullock, G. (1994). *The Economics of Non-Human Societies*. Tuscon, AZ: Pallas Press.
- Turchin, P. (2016). *Ultra Society: How 10,000 Years of War Made Humans the Greatest Cooperators on Earth*. Storrs, CT: Beresta Books.

- Tyler, T. R. (2006). *Why People Obey the Law, 2nd Edn.* Princeton, NJ; Oxford: Princeton University Press. doi: 10.1515/9781400828609
- Tyler, T. R. (2017). "Value-driven behavior and the law," in *The Oxford Handbook of Law and Economics, Volume I: Methodology and Concepts*, ed. F. Parisi (Oxford; New York, NY: Oxford University Press), 402–421. doi: 10.1093/oxfordhb/9780199684267.013.030
- Veblen, T. B. (1898). Why is economics not an evolutionary science? *Quart. J. Econ.* 12, 373–397. doi: 10.2307/1882952
- Veblen, T. B. (1899). *The Theory of the Leisure Class: An Economic Study in the Evolution of Institutions.* New York, NY: Macmillan.
- Vromen, J. J. (2010). On the surprising finding that expected utility is literally computed in the brain. *J. Econ. Methodol.* 17, 17–36. doi: 10.1080/13501780903528945
- Wegner, D. M. (2002). *The Illusion of Conscious Will.* Cambridge, MA: MIT Press. doi: 10.7551/mitpress/3650.001.0001
- Wegner, D. M. (2003). The minds best trick: how we experience conscious will. *Trends Cogn. Sci.* 7, 65–69. doi: 10.1016/S1364-6613(03)00002-0
- Wegner, D. M., and Wheatley, T. (1999). Apparent mental causation: sources of the experience of the will. *Am. Psychol.* 54, 480–492. doi: 10.1037/0003-066X.54.7.480
- Williams, G. C. (1966). *Adaptation and Natural Selection.* Princeton, NJ: Princeton University Press.
- Wilson, B. J. (2010). Social preferences aren't preferences. *J. Econ. Behav. Organiz.* 73, 77–82. doi: 10.1016/j.jebo.2008.09.013
- Wilson, D. S. (2015). *Does Altruism Exist? Culture, Genes, and the Welfare of Others.* New Haven, CT: Yale University Press.
- Winter, S. G. Jr. (1964). Economic "natural selection" and the theory of the firm. *Yale Econ. Essays* 4, 225–272.
- Zamagni, S. (1995). *The Economics of Altruism.* Aldershot: Edward Elgar.