

Effective Altruism and the Reified Mind

John Sanbonmatsu

The ideas of the ruling class are in every epoch the ruling ideas, i.e. the class which is the ruling *material* force of society, is at the same time its ruling *intellectual* force. . . . The ruling ideas are [thus] nothing more than the ideal expression of . . . the dominant material relationships grasped as ideas.

—Marx and Engels, *The German Ideology*

One of the most striking things about the Effective Altruism (EA) movement is its complacent view of itself. Innocent of any social origins, so its advocates suppose, the doctrine is said to represent no interest or constituency beyond that of *reason*, and to compass no ambition larger than the general happiness or well-being of all sentient life on earth. Having thus arrogated to itself the twin estates of reason and happiness, Effective Altruism presents itself to the world as the humble bearer of an ultimate truth, a solution to the great problems of history and society. Confounding its many critics, EA thus cheerfully marches on, extending its imperial reach over global philanthropy and social movement discourse alike.

We know, however, that unequal social conditions produce an unequal circulation of ideas, enabling some intellectual positions to assume preeminence over others, not because of their truth value, but because of their degree of compatibility with ascendant structures of the dominant economic system. In the early modern

era, we find Descartes's conception of animals as machines modeled on early capitalist relations and mechanization, while in the late nineteenth century we find Charles Darwin projecting prevalent laissez-faire views of society as a natural competition between individuals onto relations between species. Today, similarly, we find Effective Altruism formally mirroring the objective structures of late capitalism. The fact that EA has won the support of powerful billionaires is but one indication of how smoothly its ideas fit the status quo. The movement has indeed become virtually indistinguishable from its financial network of wealthy supporters, who now include some of the richest, most powerful people on earth.¹

Advocates of EA are nothing if not bullish about the virtues of the free market. "Effective altruists are usually not radicals or revolutionaries," explains Robert Wiblin, the director of research for 80,000 Hours and the former executive director of the Centre for Effective Altruism, because "sudden dramatic changes in society usually lead to worse outcomes than gradual evolutionary improvements." Wiblin (2015) admits that he "personally favor[s] maintaining and improving mostly market-driven economies," only because capitalism happens to be the most efficient mechanism for doing good. On this telling, EA has only an *accidental* relationship to capitalism. However, once we examine the matter closely, we find extensive homologies between capitalist institutions and norms, on the one hand, and the epistemic and normative structures of Effective Altruism, on the other.² EA can in fact be seen as a symptom of *reification*—the process under advanced capitalism by which thought and culture come to resemble the commodity form.

¹ The web of relations behind the Open Philanthropy Project is indicative: essentially a nonprofit slush fund, the OPP was created by Dustin Moskovitz and his wife, Cari Tuna, a former reporter for the *Wall Street Journal*. (Moskovitz, a cofounder of Facebook, is the youngest self-made billionaire in history, according to *Forbes*.) Open Philanthropy was in turn "incubated as a partnership between Cari and Dustin's foundation, Good Ventures, and GiveWell" (Open Philanthropy, n.d.). GiveWell's main funding, in turn, comes from the Global Health and Development Fund, which is managed by Elie Hassenfeld—a former hedge fund manager and the cofounder of GiveWell.

² My analysis broadly follows Max Weber's approach to the interdependency of economic and religious phenomena in *The Protestant Ethic and the Spirit of Capitalism*.

Reification and the Commodity

To understand reification in its simplest form, we might begin with the ancient legend of King Midas, to whom Dionysus gives the power of turning everything he touches into gold. Alas, as the king quickly learns, such a power is inimical to life: Midas cannot eat or drink, because his food turns to gold at his touch, and when he embraces his daughter, she too is transformed into lifeless gold. *Reification*—from the Latin *res*, for “thing”—operates in a similar way, turning living processes into “dead” things. But while Midas was a fictional king, *reification* is a real historical process in which capitalism progressively strips human culture and consciousness of their qualitative features.

Georg Lukács developed the theory of reification in *History and Class Consciousness* (1923; reprinted 1971). In his analysis of commodity fetishism, Marx had shown that while past civilizations produced goods for a variety of symbolic and cultural purposes, capitalism organizes human labor instead around the production of goods solely for their “exchange value,” enabling a dominant class to accumulate profit. In this system, all produced goods are treated as abstractly “equivalent” to one another, purely as *quantities*. The labor of human beings, too, is treated abstractly—as a commodity to be bought and sold on a market. With mechanized production and the scientific management of labor, workers now get treated as mere interchangeable parts in a machinery of accumulation. Capital’s need to coordinate the minute activities of workers fragments the laborer’s activity, in both time and space. Industrialization and urbanization uprooted human beings from the land and from the communal rituals of agrarian life. Meanwhile, the supremacy of standardized time—the mechanical clock counting out hours, minutes, and seconds for coordinating labor—stripped them too of any organic connection to the rhythms of nature.

Lukács’s insight was to see how this fragmented commodity process came more and more to obliterate the “human” dimensions

of our lives, reducing society to a general scheme of calculability and “rationalization” (the imposition of formal bureaucratic controls over society). Because capitalism is a “unified economic structure,” it correspondingly generates too a “unified structure of consciousness” (1971, 100). The *objective* needs of capital require “the commodity structure to penetrate society in all its aspects and to remold it in its own image” (85), causing “the structure of reification [to sink] more deeply, more fatefully and more definitively into the consciousness of man” (93). Reification is thus the means by “which every phenomenon—independently of its real and material distinctiveness” is “subjected to an exact calculus” (129). All domains of knowledge and experience are “subjected to an increasingly formal and standardized treatment in which there is an ever-increasing remoteness from the qualitative and material essence of the ‘things’ to which bureaucratic activity pertains” (99). In this way, reification comes “to cover the whole surface of manifest phenomena,” including the sciences, economics, journalism, the legal system, philosophy (208). It even “invades the realm of ethics” where, “[f]ar from weakening the reified structure of consciousness,” it “actually strengthens it” (99).³

Because reification is not a “thing,” but rather a set of cultural tendencies complexly related to the economic and technological system, we can recognize its presence only through its symptomology. The following are typical of the phenomenon:

- Calculability, or quantitative measurement, held as the supreme basis of human understanding—i.e., “the demand that mathematical and rational categories should be applied to all phenomena” (Lukács 1971, 113).

³ As with capitalist development as such, reification is an uneven process—it does not extend its influence over all sectors of culture at once, nor to equal local effect. Reification is more advanced in some regions of culture and consciousness more than others, and is likewise more strongly resisted in some places than others.

- Denigration of the qualitative aspects of the human personality, such as intuition, empathy, community, love, etc.
- A naive empiricism rooted in the fragmentation of knowledge, such that objects of cognition are viewed as discrete facts, without reference to complex social geographies.
- A method of analysis that renders complex social problems in a purely formalistic way, without deeper theoretical analysis.
- The machine, or machinic logic, treated as the emblem of true rationality.
- Ahistoricism and homogeneous temporality—time shorn of its qualitative, “merely human” dimensions.
- The modeling of life on the commodity form, such that individuals are represented as fungible, interchangeable units whose lives and deaths can be swapped out for one another—in much the same way that integers may be “swapped out” in a mathematical operation.
- A conception of human agency centered around “the individual, egoistic bourgeois isolated artificially by capitalism” (Lukács 1971, 135). Correspondingly, a voluntarist conception of social change that nonetheless affirms a purely aggregative account of persons in society.

It is symptomatic of reification that these highly distorted views of reason and social being are in turn *hidden from the reified mind itself*. The latter remains oblivious to its own social determinations—i.e., to the totality of social relations that together constitute the epistemic ground beneath its own feet. The problem is not one of simple ignorance, but rather of a pervasive bad faith that impels the reified mind to obscure the truth of its own complicity in power and domination.

All of these symptoms of reification, to varying degrees, are readily observed in the discourses of Effective Altruism.

Effective Altruism as Reified Thought

A widely circulated 2018 TED Talk by William MacAskill, the Oxford philosopher who has become the leading advocate of the EA movement, demonstrates several features of reification, including calculability, elimination of the qualitative dimensions of experience, ahistoricism, and the fungibility of life. In his talk, MacAskill stands before a giant animated graph representing the totality of human civilization through time. Along the X axis is a timeline of the species, beginning 200,000 years ago; along the Y axis is GDP per capita, measured in constant US dollars. “This is a graph,” MacAskill explains, “that represents the economic history of human civilization.” As MacAskill sets the timeline in motion, we see the centuries along the X axis swiftly disappear off the left side of the screen. Along the Y axis, however, nothing changes—economic growth is flatlined for 2,000 centuries. “There’s not very much going on there, is there?” MacAskill quips. “For the vast majority of human history,” he continues, “pretty much everyone lived on the equivalent of one dollar per day, and not much changed. But then something extraordinary happened: the Scientific and Industrial Revolutions. And the basically flat graph you just saw? Transforms into *this*.” Suddenly, the line along the Y axis takes a 90-degree turn upward. “What this graph means,” MacAskill explains, “is that, in terms of the power to change the world, we live in an unprecedented time in human history” (MacAskill 2018).

In reality, human civilization is so richly manifold, composed of such irreducibly diverse forms of embodied cultural experiences that it is not possible to generalize in comparative terms about life in past epochs. With a touch of his remote, however, MacAskill obliterates all traces of the qualitative dimensions of the human experiment. In place of everything that has given human life meaning and purpose, joy and pain, for countless generations—art, science, philosophy, religion, government, social struggle,

tradition, rituals of communal life—MacAskill substitutes a single quantitative metric—*per capita GDP*. The four-dimensional nature of our manifold existence as a species is thus collapsed into the two-dimensional Cartesian grid of a PowerPoint slide.⁴ Recall Lukács: “[E]very phenomenon—independently of its real and material distinctiveness” must be “subjected to an exact calculus” (1971, 129).⁵

Other advocates of EA, too, assume that the most “rational” approach to social problems and aggregate human suffering is to assume that they can be stripped of their qualitative features and represented as discrete mathematical units, as QALYS (quality-adjusted life years) or DALYS (disability-adjusted life years). Examining the relative utility value of becoming a physician in the developing world versus becoming one in the overdeveloped world, Benjamin Todd comments: “Using a standard conversion rate (used by the World Bank among other institutions) of 30 extra years of healthy life to one ‘life saved,’ 140 years of healthy life is equivalent to 5 lives saved.” Some careers are therefore better than others at maximizing outcomes—and we can calculate the latter using probability: “For instance, a 90 percent chance of helping 100 people is roughly equivalent to a 100 percent chance of helping 90 people” (Todd 2017). Reification is a “universal mathematics” for “calculating the effects of actions and of rationally imposing modes of action” (Lukács 1971, 109) in which human activity is not to “go beyond the correct calculation of the possible outcome of the sequence of events” (117). Effective Altruists continually revert to economic and mathematical terms to represent social problems, weighing philanthropic “investments” against “diminishing returns.” As Ayeya Cotra (2017), a senior researcher at Effective Altruism, says: “When we’re trying to calculate importance, it’s

⁴ PowerPoint is itself a significant vehicle of reification—see Tufte 2003.

⁵ Only in this way can the reified mind then “predict with ever greater precision all the results to be achieved” (88).

crucial to do the math . . . to figure out how many people a problem affects, to figure out how badly it affects them.” Cotra’s PowerPoint slide emphatically sums the matter up: “IMPORTANCE = SCALE x SEVERITY”—“ACTUALLY DO THE MATH.”

As others have observed, the self-understanding of Effective Altruists as impartial, “evidence-based” vessels of reason mirrors the *Weltanschauung* of a technocratic managerial elite for whom all phenomena can in fact be reduced to a balance statement, with losses on one side and gains on the other. In both cases, “control of reality” is to be effected through “the objectively correct contemplation” of “the abstract combinations of . . . relations and proportions” of assumed fact (Lukács 1971, 129). Inevitably, this top-down worldview leads to the over-valorization of billionaires and financiers in EA discourse, and a corresponding under-valorization of grass-roots activists and radicals. To the extent that EA can be described as a social movement, it is in fact a movement not of struggling social workers, English teachers, or iron workers, but of wealthy (mostly white and male) capitalists, analytic moral philosophers at elite institutions, and, significantly, technologists.

Technologists are frequently cited in EA’s devotional accounts of the white male entrepreneur as the savior of society.⁶ In one EA presentation, Cotra compares indiscriminate philanthropy to the missed opportunity of venture capitalists to invest in Microsoft in its early years. Displaying a photograph of Bill Gates and other nerdy young men in the late-1970s, Cotra asks, “Would you have invested in them? Most people didn’t, and now they’re worth \$290 billion. The key to being a good investor, and to being a good altruist, is to dig past first impressions and actually do the research so you’re more likely to be the one who makes the bet that pays off” (Cotra 2017). That Microsoft is a huge corporate polluter that boasts of its billions of dollars in contracts with the Pentagon, as well as one of

⁶ 80,000 Hours particularly recommends that idealists pursue careers in quantitative hedge-fund trading, management consulting, and technology start-ups.

the planet's leading super-spreaders of reification in daily life, including in public education, goes unremarked by the presenter.⁷ A similarly worshipful attitude toward technology entrepreneurs can be seen in other sectors of EA—as in the lobbying efforts of the Good Food Institute to promote cellular or synthesized meats (Sanbonmatsu et al. 2020).

As Lukács observed, “machinic” processes are central to reification; Herbert Marcuse later described “technological rationality” as “reification in its most mature and effective form” (Marcuse 2002, 172). In this context, the belief of elites that they can control and predict behavior in civil society is but an extension of the fragmented, mechanized labor process itself—i.e., the “structural analogue to the behavior of the worker vis-à-vis the machine he serves and observes”:

The distinction between a worker faced with a particular machine, the entrepreneur faced with a given type of mechanical development, the technologist faced with the state of science and the profitability of its application to technology, is purely quantitative; it does not directly entail *any qualitative difference in the structure of consciousness*. (Lukács 1971, 98; original emphasis)

Technologists view the world as an aggregate of resources to be manipulated and rearranged at will, such that “the principle of rational mechanization and calculability must embrace every aspect of life” (Lukács 1971, 91). It is not surprising, then, that Effective Altruists should frequently lionize Alan Turing, the father of the computer, nor that many of them, including Cotra, have professional backgrounds in computer science. For if mathematics is the software of reification, computerization is its literal hardware—the technological medium through which reified logics have come

⁷ The Bill and Melinda Gates Foundation has imposed a reified pedagogy, through computerization, on millions of public school students (Stecher, Holtzman, et al., 2018).

to penetrate every aspect of human consciousness and daily life, through video games and Fitbits to online shopping, pornography, and the manipulation of elections by companies like Cambridge Analytica. Few aspects of our experiences today escape mediation by computer algorithms. If reification is, as Lukács described it, a “dehumanized and dehumanizing” process in which “the personality can do no more than look on helplessly while its own existence is reduced to an isolated particle and fed into an alien system” (1971, 90, 92), then computerization is the ideal form of this instrumentality, the instantiation of a narrow conception of reason purified of “contamination” by the body or its feelings, such as love, desire, passion, empathy, suffering, etc.

An expression of the mathematical mind, computing is admired in EA as the paradigmatic model of consciousness itself, as such, to such a degree that its proponents seem to model their *own* subjectivity on the disembodied logic of the computer. “I don’t know about you,” confesses Cotra, “but I’m a bleeding heart. If I were to just make up numbers for how important each [philanthropic] cause was, everything would be an 11 on a scale from 1 to 10. But there’s going to be a world of difference between two causes that both seem like urgent life and death situations.” In the interests of “fairness,” then, we must “ruthlessly prioritize among causes” (Cotra 2017). Other Effective Altruists, too, caution against “choosing with the heart” or “going with our gut” when trying to promote the good. The movement is consequently hostile toward sentiment in general and empathy in particular. (Paul Bloom’s book, *Against Empathy: The Case for Rational Compassion*, is widely admired in Effective Altruism circles [Elmore 2016].) The “rational” altruist tells us to place our trust neither in moral intuition nor in elemental fellow-feeling, but rather in the high priests who keep the numbers—bankers, policy wonks, economists, elite academics, and, especially, AI researchers.

In the 1960s, Norbert Wiener (1968) and other computer scientists fantasized about achieving perfect cybernetic control

over the totality of human life. Since then, technologists attached to the national security state and corporate capital have sought to remake society in the image of mathematical machines. From their vantage point, there is no problem that cannot in theory be solved using a form of instrumental reason alienated from nature, the body, and the social. Effective Altruists too have made artificial intelligence (AI) central to their technocratic vision of mastery over social problems, with 80,000 Hours noting that “the next few decades might see the development of powerful machine learning algorithms with the potential to transform society” (80,000 Hours). While Effective Altruists warn that AI poses an “existential threat” to our species, suggesting that the emergence of a “superintelligence” could threaten human autonomy, they nonetheless embrace AI as a way to make the world “better,” counseling budding altruists to pursue careers in “top AI labs.” Unfortunately, however, since EA is unable to comprehend the *social* basis of technology in the structure of domination, its advocates fail to recognize its role in concentrating state and corporate power. 80,000 Hours thus encourages idealists wanting to change the world to build careers within the apparatus of the US national security state, by joining the Office of the Secretary of State, the National Security Council, or DARPA, the Pentagon’s cutting-edge research arm (“The Highest Impact,” 80,000 Hours)—despite the fact that the US spends nearly a trillion dollars annually on war-making, indiscriminately bombs civilians, props up dictatorships, and imposes unequal terms of trade on Third World economies. Somehow, on its path to “doing good,” EA has wound up promoting radical evil.

Irrationality and Crisis

Under reification, “quantity alone determines everything,” and time itself “sheds its qualitative, variable, flowing nature, [and] . . .

freezes into an exactly delimited, quantifiable continuum filled with quantifiable ‘things’” (Lukács 1971, 89–90). This formulation, amounting to what Walter Benjamin (2019) termed “empty, homogeneous time,” corresponds more or less exactly to the temporality depicted in MacAskill’s TED talk—his collapse of human species history into a timeline of per capita GDP. Not content to homogenize the past, however, MacAskill in the same talk projects “empty, homogeneous time” onto the future, too. EA’s proponents are in fact never more eloquent or ecstatic than when speaking of humans *who do not yet exist*, whose lives and interests they nonetheless imbue with greater moral importance than the *merely existing* humans and nonhuman animals of the present. By colonizing other planets, MacAskill thus maintains, *Homo sapiens* might live for “billions” more years, while EA advocate Toby Ord, in his bestselling book *The Precipice*, similarly invites the reader to imagine the “millions of generations” of future humans yet to come—provided only that we first dispatch the “existential threats” facing our species. Given the imminent collapse of the earth’s ecosystem, such views—which characterize existence only in terms of *quantities* of experience—are not so much optimistic as dissociative.

This homogeneous rendering of time finds its complement in the occlusion of historical fact—as when MacAskill credits the growth of GDP to “the Scientific and Industrial Revolutions,” rather than to the birth of capitalism. That MacAskill fails even to mention capitalism—the chief structuring principle of human economic and social life for the last five hundred years—is hardly an accident: only by mystifying the social origins of economic growth can he sell his cheerful vision of transhistorical progress. For to admit where all this miraculous wealth came from—*viz.*, the violent appropriation of the resources, lives, and labor of countless millions of humans and nonhumans—would otherwise require him to confront such horrors as the Atlantic slave trade, the genocide of Indigenous peoples in the Americas, Australia, and New Zealand, the destruction of the great forests of Europe, and the extermination of billions of

land and sea animals to satisfy Europe's insatiable markets for fur, fish, meat, and whale oil.

It is ironic, in this connection, that EA should pride itself on being "evidence-based" when its naive rejection of historicism and critical theory renders it *anti-empirical* in orientation. Ostensibly, the "principle of rationalization" enables the knower to "to predict with ever greater precision all the results to be achieved" (Lukács 1971, 88). In reality, however, a chasm opens up between the form and content of knowledge—i.e., between the conceptual apparatus of the "knower" and the actual content of social life. Trapped within a reified system with which it "[harmonizes] its own structure" of thought (95), the reified mind is only able to "grasp what it itself has created" (121–122). It "surrenders to the immediate facts," and in so doing "repels recognition of the factors behind the facts, and thus repels recognition of the facts, and of their historical content" (Marcuse 2002, 101). Effective Altruism's empirical inadequacy is for this reason incurable, since the "facts" that it posits are shorn of their wider sociohistorical context and significance. Because reification leads "to the destruction of every image of the whole" (Lukács 1971, 103)—occlusion of the totality of social relations—the Effective Altruist is chronically "unable to grasp the meaning of the overall process as it really is," the "'organic' unity of phenomena" (182, 188). This renders EA incapable of perceiving the patterned forces in society that lead to harm.

Though MacAskill, Peter Singer, and other Effective Altruists make much of the "good" that the rich effect in giving away portions of their fortunes in "effective" ways, the philanthropy of the rich recedes into insignificance alongside the global destruction wrought by concentrated wealth. Since the signing of the Paris Agreement on climate change in 2016, for example, "banks have facilitated almost \$4 trillion of financing for fossil fuel companies, including \$459 billion worth of bonds and loans for oil, gas and coal companies" in 2021 alone (Gelles 2021), a figure that is an

order of magnitude greater than all US philanthropic giving in the same period. Some of the same corporate leaders praised by Effective Altruists for having committed themselves to a net-zero carbon future have meanwhile resisted policy changes that could threaten corporate bottom lines. At the COP26 UN Climate Change Conference in Glasgow in 2021, for example, Jamie Dimon, the CEO of JP Morgan Chase, who has championed climate reform, pushed back against the demands of radical activists at the conference, warning that it was important for banks “to keep funding conventional energy production.” “You’re not going to get rid of oil and gas consumption tomorrow,” he told reporters (Gelles 2021). Financiers at the summit furthermore blamed the divestment movement for energy shortages and soaring energy prices, with Laurence D. Fink of BlackRock—a leading corporate figure in “conscious capitalism” (Currie 2020)—warning that transitioning too quickly away from oil would hurt emerging economies.

The reason Effective Altruists are unable to “connect the dots” between the capitalist system and its manifest consequences is that their “philosophic critique finds itself blocked by the reality from which it dissociates itself” (Marcuse 2002, 139). Doomed to mistake its own “rational and formalistic mode of cognition” for “the only possible way of apprehending reality” (Lukács 1971, 104–105, 121), EA remains helpless before the complex mediations of culture, society, and economy, unaware “that the world lying beyond its confines, and in particular . . . its own underlying reality lies, methodologically and in principle, beyond its grasp” (104). If society really did consist merely of quantifiable facts, then EA’s faith in dispassionate reason and calculation might be justified. Alas, society does not resemble the rational scheme that effective altruists attribute to it, leaving the latter blind to the “irrationality of the total process” (Lukács 1971, 102). Within EA’s cramped intellectual rooms, there is no space for Marx or Freud, or for feminism,

critical race theory, or any other historicist framework that would enable it to comprehend the social origins of, say, authoritarian populism, male violence against women, or the destruction of animals and nature. Such phenomena simply “do not compute” within EA’s own mathematicized schema, leaving the “reified mind . . . unable to perceive a pattern in this ‘chaos’” (Lukács 1971, 105). As a consequence, the movement can take aim only at the *secondary effects of the primary phenomena*. In his TED talk, MacAskill thus misidentifies the biggest problems today as global health, factory farming, and existential threats (chiefly, nuclear war, meteor strikes, and AI “singularities”). However, the global poor suffer from adverse health outcomes because of *capitalist social relations* (i.e., from a coercive division of labor rooted in exploitation and domination); the suffering of animals stems not from “factory farming,” but from long-standing patterns of human, patriarchal domination, on the one hand, and capitalist accumulation, on the other; and though we may have good reason to worry about accidental nuclear war and stray asteroids, we face more urgent concerns today—including, and above all, the mass extinction crisis. (The latter, though by far the worst catastrophe to befall terrestrial life in 66 million years, goes strangely unmentioned by MacAskill, both in his TED talk and in *Doing Good Better*, his bestselling book.)

An inability to comprehend “the phenomenon of crisis” (Lukács 1971, 105) is thus itself one of the symptoms of the reified mind. If Effective Altruists have failed to recognize the true scale of the catastrophe, or to grasp its origins, it is because today’s global crisis—the destruction of the ecological order and the breakdown of the economic, social, and political structures that have long organized human life—is rooted in fundamentally irrational social structures, institutions, and norms of which Effective Altruists can form no definite idea. As a consequence, Effective Altruists will no doubt continue to see hopeful signs of incremental, quantitative progress in specific areas of policy—e.g., in extreme poverty or malaria reduction—right up to the moment when the entire

system collapses, leaving billions to starve to death and all animal life obliterated.

Subverting Praxis and Mystifying Social Change

The evidence suggests that EA comprehends reality only in its outermost form—in the realm of appearances or immediacy (i.e., not in its fundamental character). As such, it is unable to envision a society meaningfully different than the one we now have, and so ends up affirming a conservative politics that takes existing social arrangements for granted. (As Robert Wiblin [2015] says, “We don’t want to burn the existing system to the ground,” only “to make enduring improvements to national and international systems to ensure [that] the future is better than the past.”) Such “operational rationality,” as Marcuse termed it, seeks to improve the mechanisms of repression and control, without, however, questioning their “timeless” character. Since the “reified world appears . . . as the only possible world, the only conceptually accessible, comprehensible world vouchsafed to us humans” (Lukács 1971, 110), reality shrinks to mere “facticity,” assuming the appearance of a fixed social order with “the patina of an eternal law of nature or a cultural value enduring for all time” (157). Forms of collective action and dissent that cannot be quantified are meanwhile viewed either as irrelevant or as a threat to rational planning.

The inability of Effective Altruists to picture a meaningfully different world helps explain their contempt for grass-roots activism, radicalism, and small-scale nonprofits. If existing institutions and norms are basically the right ones, and societal problems are a matter simply of reallocating resources, then attempts to disrupt or unsettle the status quo are rightly to be viewed skeptically. However, few of EA’s own descriptions of moral life, human behavior, or history correspond to the observable features of reality.

This is especially true of the doctrine's representation of the history and phenomenology of collective action, which it falsifies. EA's claim that change occurs as the aggregate result of the rational, "evidence-based" choices of dispassionate individuals fails to comport with the history of social change, which is effected not so much through incremental adjustments as by impassioned social struggles with the force to shatter an existing status quo. Consider the following cases:

- To win voting rights, British women march in the streets, go on hunger strikes, and firebomb the homes of government officials.
- When a police squad stages a routine raid on a gay nightclub in lower Manhattan, the club's patrons respond by violently rioting (to the surprise of themselves as much as to the officers).
- To strike a blow against racial segregation, a coalition of Black Christian churches organizes a boycott of buses in Montgomery, Alabama.
- Women hold consciousness-raising groups in their homes, to share their common experiences of oppression by men.
- A Tunisian man sets himself on fire to protest the lack of democracy in his country, sparking a pro-democracy movement of millions that sweeps across the Middle East.
- An autistic teenage girl in Sweden stops attending high school so that she can hold a sign on the steps of the parliament—to demand government action on climate change.

Effective Altruists cannot easily account for these or other signal events in the history of social movements because their mechanistic, fragmented conception of the world leaves them without a proper account of human agency and will. They are unable to offer a meaningful description of the affective experiences of human beings involved in struggles to overcome structures and institutions of power and injustice. EA's notion that human agency should be purged of passionate feelings, including empathy—a recurring

theme in utilitarian thought—furthermore mirrors a wider masculinist culture that eschews compassion and valorizes domination.⁸ As phenomenologists and feminist care ethicists have shown, however, empathy plays an indispensable role in constituting our moral objects (Donovan 2011, 77–94), and is even a “precondition” for moral performance (Vetlesen 2014). Arguably, it is our very capacity to “feel” our way into the experiences of others that makes moral life possible. Edith Stein went so far as to claim that empathy is the ground of *intersubjectivity* itself (Hamington 2018).

That Effective Altruists nevertheless persist in denying these basic facts of moral and social cognition is itself a symptom of their reified worldview. They assume the dissociated stance of the “experimenter” or “pure observer” (Lukács 1971, 131), the knower who stands over or apart from “the known.” As both Hegel and Marx noted, however, objective structures are realized or brought into being *subjectively* (i.e., though the passion, will, emotion, determination, etc., of flesh and blood human beings). Such a *dialectical* conception is foreign to EA, which conceives of society as a fixed system of “facts.” Under the mantle of a supposed pragmatism, the Effective Altruist looks at the way things “really are,” then adjusts his or her expectations and goals to suit the existing reality. The trouble is, when we set out believing and acting as though the world *already is what it is*—rather than something that can become *other than it is*—we foreclose on historical possibilities that might otherwise reveal themselves to us. “Only the man who wills something strongly,” Antonio Gramsci observed, “can identify the elements which are necessary to the realization of his will,” because “strong passions are necessary to sharpen the intellect and make intuition

⁸ The repugnance of Effective Altruists for such “feminine” sentiments as compassion mirrors the movement’s undertow of misogyny. Stijn Bruers, a leading Effective Altruist in Belgium, thus denies that women are systematically disadvantaged by men, saying that “the feminist movement’s reaction against men’s rights issues is irrational, with feminists misrepresenting a lot of men’s rights activists as rape apologists.” Bruers states that he “no longer believes in something like a patriarchal system that systematically privileges men and suppresses women” (Bruers 2017).

more penetrating” (1971, 171). Reality assumes determinate form only when we exercise our emotions, passions, intellect, and will as an organic unity, in concert with other perceiving, thinking, feeling beings.

There is nothing wrong with wanting to help minimize the suffering of others, nor with wanting to use one’s limited time and resources wisely. These are sensible and admirable sentiments. (If nothing else, the success of Effective Altruism challenges us to confront more honestly the dearth of strategic thinking on the left, and the need for movements to develop more carefully worked through, long-term plans for social struggle.) However, while consequentialist theory is of use in moral philosophy, it is inadequate and even harmful as a guide to social and political emancipation. The consequentialism of both Bentham and J. S. Mill hewed closely to the common sense of the bourgeois class of the early manufacturing period—a “free market” in thought as in international trade; the isolated, monadic individual as the basis of social life; the reduction of moral life (in Bentham’s version of the “hedonic calculus”) to quantitative measures; the supremacy of formal over substantive conceptions of freedom. Today we find these same *asocial* assumptions embedded in EA discourse as well. MacAskill’s morally repugnant call for an *increase* in the number of sweatshops in the Third World (2016, 128–132) is merely the artifact of a utilitarian ideology incapable of recognizing *exploitation* as a moral or social problem.⁹

Contrary to the claims of its advocates, then, EA is neither “impartial” nor politically neutral. As reified thought, EA is “anti-critical and anti-dialectical,” serving to “absorb” into its own conceptual universe “the transcendent, negative, and oppositional elements of Reason” (Marcuse 2002, 100–101). Marcuse observed

⁹ Mistaking the effect for the cause, MacAskill depicts sweatshops as the consequence of extreme poverty, rather than of a world capitalist system whose economic laws generate a perpetual need for cheap labor.

that the more we resist unfreedom, the more the dominant system appropriates our instinctual longings for emancipation and turns them against us, channeling our longings into cultural forms that serve only to strengthen the overall structure of repression. EA is but the latest entry in this dismal losing game. Trapped in the web of its own conceptual antinomies—reason vs. feeling, pragmatism vs. “idealism,” quantity vs. quality—EA is unable to identify the root of our problems or to suggest plausible means for overcoming them. One need not doubt the good intentions of individual Effective Altruists to conclude that their approach ironically preserves the very institutions that cause humans and nonhumans the most suffering. Effective Altruism is not merely unhelpful; it undermines human collective yearning for what Lukács termed an “authentic humanity, the true essence of man liberated from the false, mechanizing forms of society.”

References

- Benjamin, Walter. (2019). “Theses on the Philosophy of History.” In *Illuminations: Essays and Reflections*. New York: Mariner, 196–209.
- Bruers, S. (2017). “It’s Time to Take the Red Pill and Change Your Mind.” In *Stijn Bruers, the rational ethicist* (blog), September 2. <https://stijnbruers.wordpress.com/2017/09/02/its-time-to-take-the-red-pill-and-change-your-mind/>.
- Clean Meat Hoax. (n.d.). <https://www.cleanmeat-hoax.com/>.
- Cotra, A. (2017). “Introduction to Effective Altruism.” Address at the Effective Altruism Global X conference (EAGx Berkeley 2016); transcribed July 14, 2017. <https://forum.effectivealtruism.org/posts/5EqJozsDdHcF7dpPL/introduction-to-effective-altruism-ajeya-cotra>.
- Currie, L. (2020). “The Rise and Rise of Conscious Capitalism.” In *Meet the Leader* (blog), January 19. <https://www.meetheleader.com/blogs-1/blog-the-rise-and-rise-of-conscious-capitalism>.
- Donovan, J. (2011). “Sympathy and Interspecies Care: Toward a Unified Theory of Eco- and Animal Liberation” In *Critical Theory and Animal Liberation*, edited by John Sanbonmatsu. Lanham, MD: Rowman & Littlefield, 277–294.
- 80,000 Hours. <https://80000hours.org/career-reviews/ai-safety-researcher/>.

- Elmore, H. (2016). "We Are in Triage in Every Second of Every Day" Praises Paul Bloom's "Against Empathy: The Case for Rational Compassion." *Effective Altruism Forum*, August 6. <https://forum.effectivealtruism.org/s/YCa8BRQoxKbmf5CJb/p/vQpk3cxdAe5RX9xzo>.
- Gelles, D. (2021). "With Climate Pledges, Some Wall Street Titans Warn of Rising Prices." *New York Times*, November 4. <https://www.nytimes.com/2021/11/04/business/cop26-wall-street-pledges-fossil-fuels.html>.
- Gramsci, A. (1971). *Selections from the Prison Notebooks*. Edited by Quentin Hoare and Geoffrey Nowell Smith. New York: International Publishers.
- Hamington, M. (2004). *Embodied Care: Jane Addams, Maurice Merleau-Ponty, and Feminist Ethics*. Urbana: University of Illinois Press.
- Lukács, G. (1971). *History and Class Consciousness*. Cambridge, MA: MIT Press.
- Marcuse, H. (2002). *One-Dimensional Man*. 2nd ed. New York: Routledge.
- MacAskill, W. (2018). "What Are the Most Important Moral Problems of Our Time?" TED Talk, April 2018. https://www.ted.com/talks/will_macaskill_what_are_the_most_important_moral_problems_of_our_time?language=en#t-14221.
- MacAskill, W. (2016). *Doing Good Better*. New York: Penguin.
- Marx, K., and F. Engels. (1970). *The German Ideology*. Edited by C. J. Arthur. New York: International Publishers.
- Open Philanthropy. (n.d.). "Press Kit." <https://www.openphilanthropy.org/about/press-kit>.
- Stecher, B. M., D. J. Holtzman, et al. (2018). *Improving Teaching Effectiveness: Final Report: The Intensive Partnerships for Effective Teaching Through 2015-16*. Santa Monica, CA: RAND Corporation. https://www.rand.org/pubs/research_reports/RR2242.html.
- Todd, B. (2017). "Can One Person Make a Difference?" 80,000 Hours, April 2017. <https://80000hours.org/career-guide/can-one-person-make-a-difference/>.
- Tufte, E. (2003). *The Cognitive Style of PowerPoint*. Cheshire, CT: Graphics Press.
- Vetlesen, A. (2014). *Perception, Empathy, and Judgment: An Inquiry into the Preconditions of Moral Performance*. University Park: Pennsylvania State University Press.
- Wiblin, R. (2016). "Effective Altruists Love Systemic Change." 80,000 Hours, July 9. <https://80000hours.org/2015/07/effective-altruists-love-systemic-change/>.
- Wiener, N. (1968). *The Human Use of Human Beings: Cybernetics and Society*. London: Sphere Books.