Moral Robots

Elisabeth Huh

Ts there a best way to live? In *The Most Good You Can Do: How Effective* Elisabeth Huh is a third-year Altruism Is Changing Ideas About Living Ethically¹, moral philosopher in the College Peter Singer announces there is a key to living the most fully ethical ^{majoring in} Fundamentals. life, and it turns on a simple principle: do the most good you can. Known internationally for his 1972 essay, "Famine, Affluence, and Morality," and for his foundational work in animal rights philosophy, 1. Peter Singer. Animal Liberation, Singer has seized upon his intellectual capital in You Can Do: How this newest work to advocate for effective altruism, what he defines Effective Altruism as "a philosophy and social movement which applies evidence About Living Ethiand reason to working out the most effective ways to improve the ^{cally (New Have}Yale University world."2

Like many of Singer's works, The Most Good You Can Do is 2. Ibid., 4. polemical. Some reviewers, such as Stanford Medical Professor Walter M. Bortz, have written fervent panegyrics lauding Singer as "the voice of our collective conscience...our moral compass"³ and ^{3. Bortz, Walter} his book as "a miracle grow tonic"⁴ for your brain. Others, such as 100: Peter Singer's John Gray of The New Yorker, deems the effective altruism movement Huffington Post, one of the "follies of philosophy,"⁵ and Judith Lichtenberg of The New ^{June 2, 2015.} Republic expresses concern for the way his arguments "distort human 4. Bortz, Walter psychology."⁶ The crux of most of this controversy circles around his M. "Dare to Be 100: Book Review: longstanding philosophical claim: that we must ignore feelings and Singer, The Most personal preferences to make correct moral decisions and live the Good You Can Do." Huffpost Books. most ethical lives. This argument is evidenced from Singer's praise of Matt Wage, a former student whom he upholds as a paragon of effective altruism. After his acceptance to Oxford for postgraduate

The Most Good Is Changing Ideas cally (New Haven: Press, 2015).

M. "Dare to Be Mighty Voice."

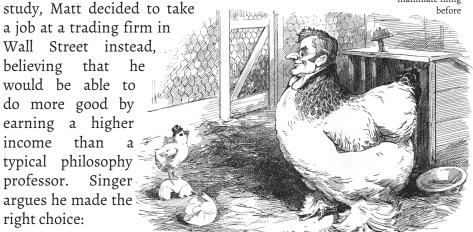
I never saw so much expression in an

inanimate thing

5. Grav. John. "How & How Review of Peter Good You Can Do: believing How Effective Altru-Ideas About Living Singer, New York Review of Books, May 21, 2015. income

Judith. "Peter Singer's Extremely Altruistic Heirs." The New Republic, November 30, 2015.

Not to Be Good." a job at a trading firm in Singer's The Most Wall Street instead, he that ism Is Changing would be able to *Ethically, by Peter* do more good by earning a higher than typical philosophy 6. Lichtenberg, professor. Singer argues he made the right choice:



One year after graduating, Matt was donating a six-figure sum roughly half his annual earnings—to highly effective charities. He was on the way to saving a hundred lives, not over his entire career but within the first year or two of his working life and every year thereafter.7

7. Singer, 6.

Singer adds later in his book that we ought to admire effective altruists not only for the positive consequences of their actions, but for their integrity. He contrasts Matt's decision with that of another hypothetical student who chose instead to accept an offer for graduate school, to write a thesis on Beowulf, and to become a professor of medieval literature. He explains that if this student also held the conviction that we ought always to try to do the most good we can, her selfish pursuit of her own passions would have cost her her ethical integrity.⁸ Singer argues that the most ethical 8. Ibid.,, 49. individuals substitute reason in the place of emotion. While Singer does believe individuals should take advantage of their unique skills and talents, he urges all those who wish to do good in the world to rationally consider the probability of saving lives through paths such as academia as long-shots.9

9. Ibid., 66.

Singer presents the career strategy of "earning to give," on the other hand, as a safe bet and as an objective good. That is, as long as we give to the most effective charities. Singer praises how effective altruists resist giving to "whatever cause tugs most strongly at their heartstrings" to maximize their positive impact.

He warns against donating to highly ineffective but emotionally compelling charities such as the Make-A-Wish Foundation, which fulfills the wishes of children suffering from life-threatening illnesses like Leukemia. He laments how a donation of \$7500 spent to fulfill one child's wish "to be Batkid" for a day could have been put to far better use by the Against Malaria Foundation to save the lives of three children. As a utilitarian, Singer's philosophical arguments run through a moral calculus that takes on what nineteenth-century philosopher Henry Sidgwick would call "the point of view of the universe."10 He distills complicated ethical choices into simple, objective consequences, weighs the newly-commensurable outputs, 10. Henry Sidgand produces what appear to be self-evident moral claims: "Saving a of Ethics (Camchild's life has to be better than fulfilling a child's wish to be Batkid." According to Singer, the emotional pleasures we derive from certain ^{2012), p. 185.} actions (such as putting a smile on a sick child's face) hinder optimal ^{11. Singer, 6.} ethical decision-making; our emotions often divert us from the path towards the good.

To be clear, I do admire Singer's decades-long crusade against all forms of suffering. I admire his unremitting efforts to push a background of inequities into the consciousness of the rich, to force them to wrestle with the weight of privilege, and to question whether they can fully digest their expensive purchasing decisions while seated at a global table serving others suffering and starvation. I also firmly agree that we all ought to broaden our sphere of moral concern and do all we can to alleviate needless suffering. But Singer is no moral saint; and I feel disquieted by his cult-like following, and by a number of insidious problems in his theory of moral progress. Singer's extreme views make him an easy target for criticism, but my primary aim is not to echo frequent arguments against utilitarianism or to repeat the other common concerns of his critics. I am interested in examining the phenomenon of Peter Singer. I want to investigate the reasons for his popularity, and I will try to do so by examining the philosophy of his philosophy; I want to discern what he presupposes as the measure of moral truth and the key to societal progress and to explain how these assumptions fit within a zeitgeist that looks fondly upon individualism, scientific rigor, rationality, and empiricism. I believe that upon close inspection,

wick. The Methods University Press

Singer's philosophy runs with rather than against the current of many of our society's problematic cultural and ethical values. I worry that this course is what may lead us morally astray.

The phenomenon of Peter Singer strikes me as an ancient Platonic dialogue unfolding on a contemporary stage. Singer himself is something like a modern Socrates; he practices frugal living, probes at the public's unquestioned habits with eccentric arguments, and philosophizes like a stinging gadfly-waking his fellow citizens from the lulling comforts of moral complacency. His efforts to influence the public through reason also recall Plato's theory of human motivation, illustrated by his conception of the tripartite soul. In Book II of The Republic, Socrates explains that the root of each human desire can find its origin in one part of a threepart soul: the lowest part, epithumia, motivates the appetites and the desire for bodily pleasure; the middle part of the soul, *thumos*, desires honor and a competitive form of self-interest; and the highest part of the soul, logos, loves reason and knowledge.¹² Socrates explains ^{12. Cooper,} that the size of each of these three parts varies among individuals, Theory of Human but that the largest piece of each person's soul naturally guides Motivation." His-her into one of three social classes. The majority of the population,

John M. "Plato's tory of Philosophy Quarterly, Vol. 1 motivated by bodily appetites, become (1984): p. 15.

the moneymakers and craftsmen of the city and fulfill its basic needs; those who desire honor are best fit to serve as the guardians and warriors of the city; and philosophers, whose actions are ruled above all by reason, ought to rule.

just The most city, according to Plato, is one ruled by a philosopher king—it is a city that models itself after a perfectly ordered soul. The limited of philosophers, supply and the utter dearth of any



Still I will proudly declare that there is something queer about it.

possessing the skill or desire to rule, posed to Plato a perpetual obstacle to justice. Singer's appearance, however, seems to mark a shift in this historical paradigm. By teaching an increasing number of ethically percipient individuals to live modestly, eschew the glory of personal wealth, and orient their lives by the light of his rational moral principles, Singer manifests, in some ways, Plato's socially conscious philosopher king. Just as the ideal philosopher king would strive to guide his city to justice through his unique understanding of the good, Singer implores the public to create a more just world by adopting his representation of the impartial perspective of the universe while making moral decisions.

So is Singer our philosophical messiah, or our modern-day Socrates? Not quite. For the few, yet essential, differences between Singer and Socrates indicate some of the most questionable qualities in the newer model's philosophy. Socrates questions, Singer pontificates. Socrates insistently professes his ignorance, while Singer exudes utter confidence. Socrates is willing to question his own assumptions and employs a dialogical form of reasoning, responsive to the principles, emotions, and concerns of his interlocutors; Singer's understanding of "reason" is surprisingly narrow—divorced from all telluric concerns, and deferential to only a few absolute principles.

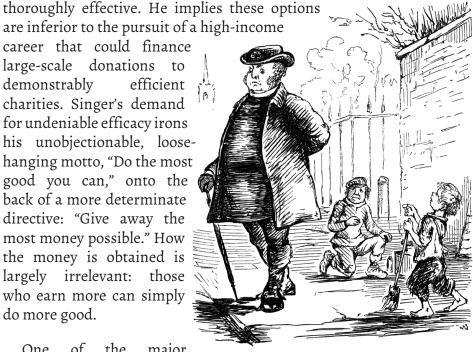
This approach appears infallible so long as these principles are truly correct. So Singer dares the polis to examine his arguments and find any reason to object to his axioms: suffering is bad, and if we can do something to prevent suffering without comparable cost to ourselves, we should. If we assent to these principles, as Singer asserts any rational person aspiring to goodness should, he then beseeches us to respect elementary math and apply its logic to his arguments in order to live the most ethical lives possible. The conclusion, he believes, is obvious: the greater number of lives we save, the more ethical we become.

Singer argues that the entire history and future of humanity's moral progress depends upon our willingness to accept his arguments. Citing Yale Psychology Professor Paul Bloom, Singer argues that it is not possible to widen our empathy indefinitely to

embrace all of humanity as our common family. He argues that instead of hoping to stretch our emotional empathy, which often leads us astray, we should instead recognize through reason that the lives of strangers hold the same value as the lives of those we love. We should numb our emotions to the images of crying children disseminated by ineffective charities and focus on identifying the most efficient means of giving. In this way, Singer argues we can subvert David Hume's famous dictum—we can enslave our emotions to our reason, and then enjoy a rational swell of satisfaction after they have served its righteous commands.

Singer also cites the research of Steven Pinker's The Better Angels of Our Nature to impute the overall decline in daily violence over the course of human history to the development of our higher reasoning capacities. Indeed, Singer depicts the most prominent adherents of effective altruism as exemplars of evolved human moral psychology, describing how they are deeply "sensitive to numbers,"¹³ how they ^{13. Singer, 78.} "talk more about the number of people they are able to help than about helping particular individuals,"¹⁴ and how they tend to have 14. Ibid., 89. expertise in areas requiring abstract reasoning, like mathematics or computing. Applauding their critical outlook towards bias and emotion, Singer upholds the effective altruists' particular mode of reasoning—predisposing them to determine the good through hard, empirical data—as a gift, aiding their recognition of deep, moral truth.

But is the key to humanity's moral progress truly impartiality and empiricism? Do these two qualities constitute the essence and pinnacle of human reason? In his book, Singer aims to convince by giving examples of effective altruist reasoning: if effective altruists have \$10,000 to donate, he states, "they would rather give it to a charity that can save a life for \$2,000 than one that can save a life for \$5,000 because they would rather save five lives than two."¹⁵ It is 15. Ibid., 78. the profit-maximizing sentiment embedded within Singer's highest moral principle, "Do the most good you can," that motivates my primary objections. Though Singer provides examples of a few other "ethical careers" people can pursue—as policymakers, organizers, advocates, researchers-he also explains that the good any individual can achieve in these fields is less predictable, and thus less



One of the major

pitfalls then, of this moral doctrine, is that it champions capital as the ultimate instrument for ethical good. By heeding the logical maxims of effective altruism, figures like Bill Gates and Warren Buffet become the altruists who live the best ethical lives simply by virtue of their higher capacity for maximally impactful financial giving. If everyone were to fully accept this logic, the result would be pernicious: our social hierarchy, already largely determined by wealth, would acquire an additional ethical justification. Those who suffer from systemic injustices, discrimination, or other societal forces of inequality would receive little aid or compassion, for the effective altruist's moral doctrine would mandate that they overlook middle-to-low classes as truly mediocre individuals whose suffering from a removed, universal perspective would be negligible. The most wealthy altruists would not only be perceived as the most powerful, but the most ethical and good. For in this system of thought, beneficence becomes the preeminent ethical virtue. Courage, creativity, compassion, and other virtues necessary for reforming systemic problems or healing complicated ailments, such as racism, xenophobia, or mental illness, become valueless atavisms.

The other issue at stake is the question of the very nature of ethical living, and whether or not ethics can be reduced to a system or to a set of determinate principles. If so, moral philosophy simply consists of their identification and application. But is it possible to discover a perfect moral decree? And even then, if we spent our entire lives obediently following its perfect code of instructionsregardless of the content of our own sentiments and beliefs and reasoning processes—could we truly become maximally ethical?

A major problem with Singer's conception of ethical living is that it requires that we mechanize ethical thinking. It turns over our ethical deliberations to the processing of a machine—programmed to maximize the reduction of suffering and to overlook the value of anything outside of this simple injunction. This is the moment, I argue, where the greatest threat of evil and wrongdoing arises: when our reason simply ministers to an unquestioned absolute. Hannah Arendt is famous for observing this phenomenon-what she called the 'banality of evil'—while describing Adolf Eichmann's 16. Hannah Ar- behavior during his trial.¹⁶ Arendt describes her horror listening Jerusalem: A Report to the listless old man defend his service to Hitler with mindless on the Bandity of Evil. (New York: stock-phrases and clichés: he did his duty, he only obeyed orders, Viking Press, he was just following the law.¹⁷ Arendt argues that this is the true terror of evil: that it comes not from sadistic, evil masterminds, but ^{17. Ibid., 27.} from the unexceptional, from the common-place, from those who become convinced to do something once and who do not stop to ask questions in their quest to fulfill their objective.

> Singer's unfaltering allegiance to his own principles leads to alarming conclusions we cannot let pass unquestioned. To demonstrate the correct application of his principles, Singer provides a scenario in his book where an effective altruist has the option to donate \$100,000 to a local art museum or to an organization restoring vision to victims of trachoma in developing countries. While I, and most, would probably agree that it is better to donate to cure the blind, Singer's own defense of this option verges on absurdity. He states:

Suppose the new museum wing will cost \$50 million, and over the fifty years of its expected useful life one million people will enjoy seeing it

endt. Eichmann in 1963).

each year, for a total of fifty million enhanced museum visits. Since you would contribute 1/500th of the cost, you could claim credit for the enhanced aesthetic experiences of one hundred thousand visitors. What if you donate to cure blindness? ...a donation of \$100,000 could be expected to restore or preserve the sight of one thousand poor people in developing countries. On the one side, then, we have enhanced aesthetic experiences for one hundred thousand museum visitors, and on the other side we have one thousand people spared fifteen years of blindness, with all the problems that that causes for poor people with no social security.¹⁸

If Singer is aware of the problems in this argument and intentionally seeks to persuade through hyperbole, then few can single him out for this political use of rhetoric. But if Singer sincerely believes he can perfectly reduce the value of a donation to an art museum to the "enhanced aesthetic experiences" of a certain amount of people, commensurate with the size of the donation, we ought to reflect on the problematic values underlying his deeply flawed reasoning. We also ought to reflect on the types of people that might find these arguments convincing.

Singer's philosophy aims to eliminate uncertainty through quantitative reasoning. His arguments, therefore, appeal especially to those who want to make absolutely correct ethical decisions and who detest uncertainty. They resonate for the increasing number of people who perceive numbers and empirical evidence as the ultimate standard-bearers for rigor, seriousness, and truth. Singer has tapped into the heart of the values of science and capitalism to create something like a new secular religion, promising that anyone can achieve maximal goodness by simply buying into his efficient brand of ethical living. His book explains whether effective altruists should have children (though they suck money away from other suffering people, they are a great investment if you also raise them to become effective altruists); he gives examples of cheap or free hobbies altruists can adopt for fun and fulfillment; and he advises how to manage a budget that settles just above the point of marginal utility.

In a world of suffering and inequality, Singer argues that those

18. Singer, 119.

who wish to live by his doctrine should consider art museums as nothing more than buildings offering "aesthetic experiences." But by contrasting the value of an "aesthetic experience" with the value of saving a life, Singer falsely represents the value of museums and overlooks the essential role that art plays in a great number of lives. What of autistic individuals who use art as their primary means of communication? Or those who sell or create art for a living? What about the millions of people whose experiences and reflections in an art museum might have helped them to perceive the world in new ways, perhaps inspiring new ideas and projects that then helped to shape the world? The fact that the impact of an art museum is not easily quantifiable does not mean that it does not have important value. Just because a smaller population may consider art essential does not mean its value is objectively negligible. The dichotomy Singer creates between "objective" physical needs and "subjective" experiences treats our need to fill our lives with subjective content as frivolous fluff. Singer asserts that from the perspective of the universe, an art museum holds no value outside of aesthetic experience, that subjective experiences are selfish and unimportant, and that anyone who disagrees with him is objectively wrong.

So I object to his logic, even if that makes me "objectively wrong," for I believe that even if we accept the salience of numbers and impartiality in areas like science, we should reject that

ethics should work in the same way. We should reject the notion that we can quantitatively determine the most serious forms of suffering through an algorithm, and ideal that moral thinking should consist of nothing more than our pursuit of this end. I wholeheartedly protest against the



This paper looks to me as if it KNEW what a vicious influence it had!

claim that we should only help those who are "objectively" the worst off, and that we are undoubtedly better off when we strain every effort to alleviate suffering through the sieves of certainty and efficiency. For if these claims are true, we should outsource our moral thinking to robots, who are truly best fit to act as perfect moral agents. If Singer is right, we should define humanity's distinct capacity for 'moral reasoning' as nothing more than the recognition of logical principles, the cognitive capabilities required for mathematics, and the thinking required to maximize efficiency.

But is Singer right? Is the key to moral progress truly greater objectivity and impartiality, and should we aspire to be more like robots? I think not. For I believe the human reason that drives moral progress does not lack individual, subjective content; this content is what imparts its human nature. I believe we should actually look to Wittgenstein—a philosopher who notoriously denies the existence of all objective, absolute, or metaphysical truths-for some of the deepest insights into how ideal moral thinking should operate. In ^{19. Ludwig} Wittgenstein, his Philosophical Investigations, Wittgenstein argues that the role of Philosophical Invesphilosophy is not to make discoveries or unveil absolute truths by E. M. Anscombe, means of pure logical rules; rather, it aims to understand meaning by ^{P. M. S. Hacker,} surveying our use of language in our everyday lives.¹⁹ The philosopher Schulte, rev. 4th Cora Diamond shows how Wittgensteinian thinking applies to Wiley-Blackwell, moral philosophy in her analysis of Plato's dialogue, Crito.²⁰ In the Crito, Socrates argues that he should not escape from prison because 20, Cora Diamond. we all ought to obey our parents and teachers, and if he escapes, he "Missing the Adventure: will be disobeying the state of Athens, which has been his parent and Reply to Martha teacher. An absolute or formulaic conception of moral reasoning Realistic Spirit: would argue that Socrates' moral argument consisted of finding a Wittgenstein, Phi-losophy, and the general moral rule or principle-it is wrong to disobey our parents Mind (Cambridge, and teachers—by making a statement of fact—that the State is his Press, 1991). teacher—and finally, by generating his moral decision—it is wrong to disobey the State-by applying fact to principle. Wittgenstein would object to this interpretation, however, for it asserts that philosophy should simply be used like math or logic to discover an answer to a problem already objectively fixed in advance.

2009).

Nussbaum," in The MA: The MIT

Diamond presents an entirely different take on this situation. She explains that Socrates' moral thinking did not simply consist

29

in identifying fixed logical truths, and that he did not merely state objective facts. She argues that Socrates employed startlingly original moral thinking in describing Athens as his "parent" and "teacher." In testing the boundaries of the words "parent" and "teacher," Socrates reframed the facts of the situation at hand to transform Crito's own moral perceptions. Applying his own interpretation of fact to a general understanding of right and wrong, Socrates also took on full, subjective responsibility for his moral decision. Socrates did not use philosophy like a lantern to discern some objectively correct course of action, pre-written on some metaphysical moral plane, he fused his own creative, subjective deliberation into his use of reason to present what he considered the right thing to do.

Socrates also exhibited sensitivity to Crito's feelings. Crito beseeched Socrates to flee from prison out of compassion for his friends and children. Socrates responded, not by denying the validity of Crito's concerns, but by acknowledging that he shared Crito's essential moral intuition—that it is wrong to treat others badly and by then showing how this same value justified his own decision: he should not escape because it would do harm to Athens. Socrates uses this point of agreement as a pivot to shift Crito's perspective, to share his own beliefs, and to explain his disagreement in terms he hopes Crito can empathize with and understand.

We cannot, and should not, seek to wholly eliminate subjective thinking or the emotions from our moral thinking, for these are essential elements of human existence that we ought not to scorn as unimportant or undesirable. The emotions manifest human values, cultivated by our experiences, constitutive of cognitive processes, and shaped by our reason. We feel grief at the loss of something we had cherished and loved; we fear that which poses a threat to our understanding of well-being; we feel anger when we sense injustice and wrongdoing. Though Singer does not go so far as to claim that effective altruists ought to completely strip away their emotions, he does insist that the emotions distract from his principle of efficiency, and that this principle is essential for optimal moral decision-making. So this is what I consider his ultimate mistake: it is his conflation of morality with efficiency, and his belief that we do not need the emotions and some acceptance of uncertainty on our

path to moral progress.

Why is this a mistake? Let's look at the legalization of gay marriage in the United States. Singer's effective altruist doctrine would have maintained that the suffering of homosexual couples was not as great as the suffering of those starving to death, and that, therefore, "maximally effective" altruists wishing to do "the most good" possible should not have considered spending any time, money, or resources fighting for some the right to some highly subjective form of emotional fulfillment.

Yet a great number of individuals "defied" reason and logic by listening to their hearts anyway. They fought for what the perspective of the universe might consider a completely selfish, emotional, and subjective form of self-betterment-and the result was a landmark sign of social progress. The right to same-sex marriage was unscripted—it was not a victory activists could have definitely predicted in advance—and its value would have been difficult to quantify (although, I suppose, some could argue married homosexual couples make more money, enjoy certain tax and welfare benefits, and raise happier children, and then total the sum of that eudaimonistic calculus). Still, the U.S. Supreme Court decision in Obergefell v. Hodges recognized the salient desires of homosexual couples, and they reached this decision through moral reasoning.

In his decision, Justice Kennedy described marriage, not as a 21. Obergefell v. strict legal relationship between a man and a union, but as a lifelong $\frac{Hodges 576 \text{ U.S. 3}}{(2015)}$ union "arising from our basic needs," "essential to our most profound hopes and aspirations," and promising "nobility and dignity to all persons."²¹ He did not discover this definition while reading through the dictates of absolute moral law, but by listening to the moral language of the human beings before him beseeching a new form of understanding. It was through listening to their arguments and 22. Obergefell v. feelings that Kennedy was able to declare, "It would misunderstand (2015). these men and women to say they disrespect the idea of marriage... They ask for equal dignity in the eyes of the law. The Constitution grants them that right."22

Hodges 576 U.S. 28

This is the end towards which I believe morality progresses—the

recognition of universal human dignity. It is the recognition that we each have unique preferences, beliefs, desires, and dreams, and that these human needs deserve universal respect. The compassionate and imaginative mental work necessary to reach this recognition and not our unthinking adherence to fixed rules or principles—is what distinguishes human reason. We ought to reflect upon the etymology of the Greek word *logos*, which refers not only to "reason," but "word," "speech," and "opinion." There is a grammar and a logical structure underlying the language of morality, but this is set into place by shared human values, needs, and experiences. Our willingness to work through this shared grammar in order to hear, debate, disagree with, and understand those who think differently from ourselves is what fuels the engines of political change and moral progress.

The reasoning of effective altruism is incredibly useful for those who possess the equivalent of a truck-load of band-aids and seek to identify where this general aid can provide the most urgent help. Yet charitable giving alone cannot heal unhealthy minds, reform insidious systemic or structural injustices, or lend a human ear to those who suffer from disrespect, discrimination, and the overall loss of their dignity. Singer arguments have garnered him impressive political capital: his mastery of the language of objectivity and empiricism has allowed him to speak to the values of a growing number of people. But it is wrong to believe that his arguments align with absolute truths. It is wrong to believe that we should deafen our ears to calls for help that do not speak with the weight of empirical urgency.

Efficiency, impartiality, and logical consistency are virtues of math and science that should not continue to carry the same valence in the sphere of ethics. We can never discover an absolute ethical system; and we do not make moral progress by spurning humanity in the pursuit of some perfection above ourselves. We use ethics and moral philosophy to try to understand how to live well, and we make continual progress by listening to our humanity—by perpetually struggling to amend our rules and systems to fit a constantly changing, eternally imperfect, and unalterably human form of existence.