

What Is Effective Altruism?

I met Matt Wage in 2009 when he took my Practical Ethics class at Princeton University. In the readings relating to global poverty and what we ought to be doing about it, he found an estimate of how much it costs to save the life of one of the millions of children who die each year from diseases that we can prevent or cure. This led Matt to calculate how many lives he could save, over his lifetime, assuming that he earned an average income and donated 10 percent of it to a highly effective organization, for example, one providing families with bednets to prevent malaria, a major killer of children. He discovered that he could, with that level of donation, save about one hundred lives. He thought to himself, “Suppose you see a burning building, and you run through the flames and kick a door open, and let one hundred people out. That would be the greatest moment in your life. And I could do as much good as that!”¹

Two years later Matt graduated. His senior thesis received the Philosophy Department’s prize for the best thesis of the year. He was accepted by the University of Oxford for postgraduate study. Many students who major in philosophy dream of an opportunity like that—I know I did—but by then Matt had done a lot of thinking about and discussing with others what career would do the most good. This led him to a very different choice: he took a job on Wall Street, working for an arbitrage trading firm. On a higher income, he

would be able to give much more, both as a percentage and in dollars, than 10 percent of a professor's income. 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.

Matt is an effective altruist. His choice of career is one of several possible ways of being an effective altruist. Effective altruists do things like the following:

- Living modestly and donating a large part of their income—often much more than the traditional tenth, or tithe—to the most effective charities;
- Researching and discussing with others which charities are the most effective or drawing on research done by other independent evaluators;
- Choosing the career in which they can earn most, not in order to be able to live affluently but so that they can do more good;
- Talking to others, in person or online, about giving, so that the idea of effective altruism will spread;
- Giving part of their body—blood, bone marrow, or even a kidney—to a stranger.

In the following chapters, we will meet people who have done these things.

What unites all these acts under the banner of effective altruism? The definition now becoming standard is “a philosophy and social movement which applies evidence and reason to working out the

most effective ways to improve the world.”² That definition says nothing about motives or about any sacrifice or cost to the effective altruist. Given that the movement has altruism as part of its name, these omissions may seem odd. Altruism is contrasted with egoism, which is concern only for oneself, but we should not think of effective altruism as requiring self-sacrifice, in the sense of something necessarily contrary to one’s own interests. If doing the most you can for others means that you are also flourishing, then that is the best possible outcome for everyone. As we shall see in chapter 9, many effective altruists deny that what they are doing is a sacrifice. Nevertheless they are altruists because their overriding concern is to do the most good they can. The fact that they find fulfillment and personal happiness in doing that does not detract from their altruism.

Psychologists who study giving behavior have noticed that some people give substantial amounts to one or two charities, while others give small amounts to many charities. Those who donate to one or two charities seek evidence about what the charity is doing and whether it is really having a positive impact. If the evidence indicates that the charity is really helping others, they make a substantial donation. Those who give small amounts to many charities are not so interested in whether what they are doing helps others—psychologists call them warm glow givers. Knowing that they are giving makes them feel good, regardless of the impact of their donation. In many cases the donation is so small—\$10 or less—that if they stopped to think, they would realize that the cost of processing the donation is likely to exceed any benefit it brings to the charity.³

In 2013, as the Christmas giving season approached, twenty thousand people gathered in San Francisco to watch a five-year-old boy dressed as Batkid ride around the city in a Batmobile with an actor dressed as Batman by his side. The pair rescued a damsel in

distress and captured the Riddler, for which they received the keys of “Gotham City” from the mayor—not an actor, he really was the mayor of San Francisco—for their role in fighting crime. The boy, Miles Scott, had been through three years of chemotherapy for leukemia, and when asked for his greatest wish, he replied, “To be Batkid.” The Make-A-Wish Foundation had made his wish come true.

Does that give you a warm glow? It gives me one, even though I know there is another side to this feel-good story. Make-A-Wish would not say how much it cost to fulfill Miles’s wish, but it did say that the average cost of making a child’s wish come true is \$7,500.⁴ Effective altruists would, like anyone else, feel emotionally drawn toward making the wishes of sick children come true, but they would also know that \$7,500 could, by protecting families from malaria, save the lives of at least three children and maybe many more. Saving a child’s life has to be better than fulfilling a child’s wish to be Batkid. If Miles’s parents had been offered that choice—Batkid for a day or a complete cure for their son’s leukemia—they surely would have chosen the cure. When more than one child’s life can be saved, the choice is even clearer. Why then do so many people give to Make-A-Wish, when they could do more good by donating to the Against Malaria Foundation, which is a highly effective provider of bednets to families in malaria-prone regions? The answer lies in part in the emotional pull of knowing that you are helping *this* child, one whose face you can see on television, rather than the unknown and unknowable children who would have died from malaria if your donation had not provided the nets under which they sleep. It also lies in part in the fact that Make-A-Wish appeals to Americans, and Miles is an American child.

Effective altruists will feel the pull of helping an identifiable child from their own nation, region, or ethnic group but will then ask themselves if that is the best thing to do. They know that saving a life

is better than making a wish come true and that saving three lives is better than saving one. So they don't give to whatever cause tugs most strongly at their heartstrings. They give to the cause that will do the most good, given the abilities, time, and money they have available.

Doing the most good is a vague idea that raises many questions. Here are a few of the more obvious ones, and some preliminary answers:

What counts as "the most good"?

Effective altruists will not all give the same answer to this question, but they do share some values. They would all agree that a world with less suffering and more happiness in it is, other things being equal, better than one with more suffering and less happiness. Most would say that a world in which people live longer is, other things being equal, better than one in which people live shorter lives. These values explain why helping people in extreme poverty is a popular cause among effective altruists. As we shall see in more detail in chapter 10, a given sum of money does much more to reduce suffering and save lives if we use it to assist people living in extreme poverty in developing countries than it would do if we gave it to most other charitable causes.

Does everyone's suffering count equally?

Effective altruists do not discount suffering because it occurs far away or in another country or afflicts people of a different race or religion. They agree that the suffering of animals counts too and generally agree that we should not give less consideration to suffering just because the victim is not a member of our species. They may differ, however, on how to weigh the type of suffering animals can experience against the type of suffering humans can experience.⁵

Does “the most good you can do” mean that it is wrong to give priority to one’s own children? Surely it can’t be wrong to put the interests of members of the family and close friends ahead of the interests of strangers?

Effective altruists can accept that one’s own children are a special responsibility, ahead of the children of strangers. There are various possible grounds for this. Most parents love their children, and it would be unrealistic to require parents to be impartial between their own children and other children. Nor would we want to discourage such bias because children thrive in close, loving families, and it is not possible to love people without having greater concern for their well-being than one has for others. In any case, while doing the most good is an important part of the life of every effective altruist, effective altruists are real people, not saints, and they don’t seek to maximize the good in every single thing they do, 24/7. As we shall see, typical effective altruists leave themselves time and resources to relax and do what they want. For most of us, being close to our children and other family members or friends is central to how we want to spend our time. Nonetheless, effective altruists recognize that there are limits to how much they should do for their children, given the greater needs of others. Effective altruists do not think their children need all the latest toys or lavish birthday parties, and they reject the widespread assumption that parents should, on their death, leave virtually everything they own to their children rather than give a substantial part of their wealth to those who can benefit much more from it.

What about other values, like justice, freedom, equality, and knowledge?

Most effective altruists think that other values are good because they are essential for the building of communities in which people can live better lives, lives free of oppression, and have greater

self-respect and freedom to do what they want as well as experience less suffering and premature death.⁶ No doubt some effective altruists hold that these values are also good for their own sake, independently of these consequences, but others do not.

Can promoting the arts be part of “the most good you can do”?

In a world that had overcome extreme poverty and other major problems that face us now, promoting the arts would be a worthy goal. In the world in which we live, however, for reasons that will be explored in chapter 11, donating to opera houses and museums isn't likely to be doing the most good you can.

How many effective altruists could there be? Can everyone practice effective altruism?

It's possible for everyone who has some spare time or money to practice effective altruism. Unfortunately, most people—even, as we shall see in chapter 11, professional philanthropy advisors—don't believe in thinking too much about the choice of causes to support. So it isn't likely everyone will become an effective altruist anytime soon. The more interesting question is whether effective altruists can become numerous enough to influence the giving culture of affluent nations. There are some promising signs that that may be starting to happen.

What if one's act reduces suffering, but to do so one must lie or harm an innocent person?

In general, effective altruists recognize that breaking moral rules against killing or seriously harming an innocent person will almost always have worse consequences than following these rules. Even thoroughgoing utilitarians, who judge actions to be right or wrong entirely on the basis of their consequences, are wary of speculative

reasoning that suggests we should violate basic human rights today for the sake of some distant future good. They know that under Lenin, Stalin, Mao, and Pol Pot, a vision of a utopian future society was used to justify unspeakable atrocities, and even today some terrorists justify their crimes by imagining they will bring about a better future. No effective altruist wants to repeat those tragedies.

Suppose I set up a factory in a developing country, paying wages that are better than local workers would otherwise earn and enough to lift them out of extreme poverty. Does that make me an effective altruist, even if I make a profit from the factory?

What are you going to do with your profits? If you decided to manufacture in the developing country in order to make it possible for people to escape extreme poverty, you will reinvest a substantial part of your profits in ways that help more people escape extreme poverty. Then you are an effective altruist. If, on the other hand, you use your profits to live as luxuriously as you can, the fact that you have benefited some of the poor is not sufficient to make you an effective altruist. There are all kinds of intermediate positions between these two extremes. Reinvesting some of your profits to help more people earn a decent income, while retaining enough to live at a much better level than your employees, puts you somewhere on the spectrum of effective altruism—you are living at least a minimally decent ethical life, even if not a perfect one.

What about giving to your college or university? You teach at Princeton University, and this book is based on lectures you gave at Yale University, thanks to the generous gift of a Yale alumnus. Do you deny that giving to such institutions counts as effective altruism?

WHAT IS EFFECTIVE ALTRUISM?

I count myself fortunate to be teaching at one of the finest educational institutions in the world. This gives me the opportunity to teach very bright, hardworking students like Matt Wage, who are likely to have a disproportionately large influence on the world. For the same reason, I was pleased to accept the invitation to give the Castle Lectures at Yale. But Princeton has an endowment, at the time of writing, of \$21 billion, and Yale's is \$23.9 billion. At the moment there are enough alumni donating to these universities to ensure that they will continue to be outstanding educational institutions, and the money you donate to one of them could probably do more good elsewhere. If effective altruism ever becomes so popular that these educational institutions are no longer able to do important research at a high level, it will be time to consider whether donating to them might once again be an effective form of altruism.⁷

This page intentionally left blank

A Movement Emerges

Effective altruism is an offspring with many parents. I can claim to be one of them because in 1972, when I was a junior lecturer at University College, Oxford, I wrote an article called “Famine, Affluence and Morality” in which I argued that, given the great suffering that occurs during famines and similar disasters, we ought to give large proportions of our income to disaster relief funds. How much? There is no logical stopping place, I suggested, until we reach the point of marginal utility—that is, the point at which by giving more, one would cause oneself and one’s family to lose as much as the recipients of one’s aid would gain. Over the next forty plus years, the essay has been widely reprinted and used by professors around the world to challenge their students’ beliefs that they are living ethically.¹

Here’s the rub: even though I argued that this is what we ought to do, I did not do it myself. When I wrote the article, my wife and I were giving away about 10 percent of our modest income (she was working as a high school teacher, earning a little more than I was). That percentage increased over the years. We are now giving away about one-third of what we earn and aiming to get to half, but that still isn’t anywhere near the point of marginal utility. One of the things that made it psychologically difficult to increase our giving was that for many years we were giving away a bigger slice of our

income than anyone we knew. No one, not even the megarich, seemed to be giving a higher proportion.

Then in 2004 the *New Yorker* published a profile of Zell Kravinsky. Kravinsky had given almost his entire \$45-million real estate fortune to charity. He did put some money into trust funds for his wife and children, but the children were attending public schools, and he and his family were living on about \$60,000 a year. He still did not think he had done enough to help others, so he arranged with a nearby hospital to donate a kidney to a stranger. The article linked my then-thirty-two-year-old essay to Kravinsky's way of living and quoted him as saying, "It seems to me crystal clear that I should be giving all my money away and donating all of my time and energy."²

By this time I was teaching at Princeton, not far from where Kravinsky lived, so I invited him to speak to one of my classes, something he has done regularly ever since. Kravinsky is a brilliant man: he has one doctorate in education and another on the poetry of John Milton. He taught at the University of Pennsylvania before turning from academic life to real estate investment, so he is at home in the university environment. Despite his interest in poetry, he puts his altruism in mathematical terms. Quoting scientific studies that show the risk of dying as a result of making a kidney donation to be only 1 in 4,000, he says that not making the donation would have meant he valued his life at 4,000 times that of a stranger, a valuation he finds totally unjustified. He even told Ian Parker, the author of the *New Yorker* profile, that the reason many people don't understand his desire to donate a kidney is that "they don't understand math."

Around the time I was reading about Kravinsky I became aware of the work of Abhijit Banerjee and Esther Duflo, professors of economics at MIT, who founded the Poverty Action Lab to carry out "social experiments"—by which they meant empirical research to

discover which interventions against poverty work and which do not. Without such evidence, Duflo points out, we are fighting poverty the way medieval doctors fought illness by applying leeches.³ Banerjee and Duflo pioneered the application of randomized controlled trials, the golden standard of the pharmaceutical industry, to specific aid projects. By 2010 researchers associated with the Poverty Action Lab—now known as the Abdul Latif Jameel Poverty Action Lab, or J-PAL—had carried out 240 experiments in 40 countries. Dean Karlan, once a student of Banerjee and Duflo and now himself a professor of economics at Yale, started Innovations for Poverty Action, a nonprofit organization to bridge the gap between academic research and the practical side of development. Innovations for Poverty Action has grown to have a staff of nine hundred and a budget of \$25 million, and the idea of randomized trials is clearly catching on.

In 2006 Holden Karnofsky and Elie Hassenfeld were in their midtwenties, working for a hedge fund in Connecticut, and earning far more than they had any desire to spend. Together with some of their colleagues, they talked about giving significant amounts to charity—but to which charity? (The Poverty Action Lab and Innovations for Poverty Action evaluate specific interventions, such as distributing bednets to protect people against malaria, but not the charitable organizations themselves, most of which have several programs.) Karnofsky, Hassenfeld, and their colleagues were used to analyzing large amounts of data in order to find sound investments. They contacted several charities and asked them what a donation would accomplish. They got lots of glossy brochures with pictures of smiling children but no data that told them what the charities were achieving and at what cost. Calling the charities and explaining what they wanted to know got them no further. One charity told them that the information they were seeking was confidential. Karnofsky

and Hassenfeld sensed a vacuum that needed to be filled. With financial support from their colleagues, they set up GiveWell, an organization that has taken the evaluation of charities to a new level. They soon found they could not run it part-time and so left the hedge fund, a move that cut their income by more than half. Their assumption is that if enough people follow the recommendations on GiveWell's website, the charities will realize that it is in their interest to be transparent and to demonstrate their effectiveness. GiveWell estimates that in 2013 more than \$17 million dollars went to its top-ranked charities as a result of those rankings. Although that is not enough to have a major impact on the charitable field as a whole, the figure has risen sharply each year since GiveWell was launched. GiveWell's existence has been critical to the development of the effective altruism movement. Now, when skeptics ask, How do I know that my donation will really help people in need? there is a good reply: If you give to one of GiveWell's top-rated charities, you can be confident that your donation will do good and be highly cost-effective.⁴

Around the time Karnofsky and Hassenfeld were setting up GiveWell, Toby Ord was studying philosophy at the University of Oxford. As an undergraduate, Ord, an Australian, had initially studied computer science and mathematics at the University of Melbourne, but he often got into arguments about ethical and political issues. When he expressed his views about poverty, his friends would retort, "If you believe that, why don't you just give most of your money to people starving in Africa?" His friends thought that this conclusion was absurd, but Toby asked himself, "If my money could help others much more than it helps me, then why not?"

Ord's growing interest in ethics led him to do a second undergraduate degree in philosophy. He did so well that he got a

scholarship to Oxford, where he wrote a doctoral thesis on how we should decide what to do. He remained interested in practical ethics, and read my article "Famine, Affluence and Morality." He began to think seriously about what he could do for people in extreme poverty. At the time he was living quite comfortably on his graduate student scholarship, which paid him £14,000 a year, a sum that put him, he noticed, in the richest 4 percent of the world's people, even after adjusting for how much further money goes in developing countries.⁵ When he graduated he would be earning more. He decided to calculate how much he would be able to give away over his lifetime, after meeting his own needs, assuming he earned a standard academic salary. His earnings, he estimated, might come to £1.5 million, or US\$2.5 million (in 2005 dollars), and of this, he thought he could donate two-thirds, that is, £1 million, or US\$1.7 million. Then he asked himself what that sum could achieve if it were donated to the most effective charities. He estimated that, while maintaining an attractive quality of life, he could donate enough to cure eighty thousand people of blindness or to save around fifty thousand years of healthy life.⁶ In other words, his donations would achieve the equivalent of saving the lives of one thousand children, each of whom would live another fifty years in good health, or of enabling five thousand people to live an extra ten healthy years. Such benefits so dramatically outweighed the small sacrifice Ord imagined he would be making that he committed himself to living on £20,000 per annum (adjusted for inflation and equivalent to US\$34,000) and giving away the rest. His wife, Bernadette Young, a physician, pledged to give away everything above £25,000 (US\$42,600). Ord subsequently lowered his own allowance to £18,000 (US\$30,600), as he found that £20,000 was more than he needed to live comfortably and even take a holiday in France or Italy.⁷

Ord wanted to share his knowledge of how easy it is to make a huge positive difference to the world. In 2009 he and Will MacAskill, another Oxford philosophy graduate student, founded Giving What We Can, an international society dedicated to eliminating poverty in the developing world. Members pledge to give at least 10 percent of their income to wherever they think it will do the most to relieve suffering in the developing world. At the time of writing, 644 people have taken the pledge, and Giving What We Can estimates that if the donors all do what they have pledged to do, \$309 million will go to the most effective charities.⁸

In addition to helping Ord launch Giving What We Can, MacAskill had an idea for another organization. Students and other young people get plenty of career advice, but none of it is directed toward the question an effective altruist would ask: What career will enable me to do the most good over my lifetime? In 2011 MacAskill and five friends founded 80,000 Hours, so named because that is roughly the number of hours people spend working in their careers. 80,000 Hours does research on which careers do the most good, offers free career coaching, and is building a global community of people seeking to change the world for the better.⁹ (Curious about the careers 80,000 Hours recommends? Wait for chapters 4 and 5.)

The term *effective altruism* was born when Giving What We Can and 80,000 Hours decided to apply for charitable status under a common umbrella organization. The umbrella organization needed a name. After tossing around some names, including High Impact Alliance and Evidence-based Charity Association, the group took a vote, and Centre for Effective Altruism was the clear winner. Effective altruism soon caught on and became the term for the entire movement.¹⁰

While these developments were taking place, I continued to write about our obligations to help people in great need. In 1999 and

A MOVEMENT EMERGES

2006 I published essays in the *New York Times Sunday Magazine*. The response to the second article was so positive that I developed it into a book, *The Life You Can Save*, which appeared in 2009. In the final chapter I suggested a progressive scale of giving, like a tax scale, where the amount you give increases as your income increases. As compared with Giving What You Can's flat 10 percent pledge, my suggestion is for a lower level of giving for average income earners but a higher one for those with high incomes. Agata Sagan, a Polish researcher and supporter of the ideas in the book, set up a website so that people can pledge online to meet the level of donation suggested for their income. So far, more than seventeen thousand people have taken this pledge. Gradually the website evolved into an organization, which really took off when I got an email from Charlie Bresler. Charlie and his wife, Diana Schott, are representative of those who, nearing the age when people traditionally retired from paid employment, are thinking about what they want to do with the next ten, twenty, or even thirty years of useful life they may still have left. As students in the sixties, Charlie and Diana had been active in the movement against the war in Vietnam and for a transformation of political life in the United States. When it became apparent that the system was more resistant to reform than they had hoped, Diana became a physician while Charlie earned a doctorate in psychology. After spending some time as a professor of psychology he stumbled—his word—into becoming president of Men's Wearhouse, a major national clothing chain. That increased his income, but in the back of their minds Charlie and Diana retained the idea that after they had raised their family they would do something to make the world a better place. Charlie read *The Life You Can Save* and decided that helping people in extreme poverty would be worthy of his time and energy. He is now the unpaid executive director of The Life You

EFFECTIVE ALTRUISM

Can Save. In 2013, the first year in which the organization was fully operational, it conservatively estimated that, on a budget of \$147,000, it had moved \$594,000 to highly effective charities, yielding a “return on investment” of more than 400 percent.¹¹