The Psychology of (In)Effective Altruism

Lucius Caviola,1,* Stefan Schubert,2 and Joshua D. Greene1,3

The most effective charities are hundreds of times more effective than typical charities, yet few donors prioritize effectiveness. Why is that? How might we increase the effectiveness of charitable giving? We review the motivational and epistemic causes of (in)effective giving. Many donors view charitable giving as a matter of personal preference, which favors decisions based on emotional appeal rather than effectiveness. In addition, while many donors are motivated to give effectively, they often have misconceptions and cognitive biases that reduce effective giving. Nearly all research on charitable giving interventions focuses on increasing donation amounts. However, to increase societal benefit, donation effectiveness is likely to be more important. This underscores the need for research on strategies to encourage effective giving.

The Importance of Effectiveness

A US$100 donation can save a person in the developing world from trachoma, a disease that causes blindness [1]. By contrast, it costs US$50 000 to train a guide dog to help a blind person in the developed world. This large difference in impact per dollar is not unusual. According to expert estimates, the most effective charities are often 100 times more effective than typical charities [2]. In the USA, people donate approximately US$450 billion [2% of gross domestic product (GDP)] each year [3], but relatively little of this supports the most effective charities – the ones that save the most lives, or improve lives the most, with each dollar (Box 1).

Most research on charitable giving focuses on the amounts that donors give [4]. However, if the societal goal of charitable giving is to improve human (or animal) well-being, it is probably more important to focus on the effectiveness of giving. This is because the variation in charity effectiveness is much larger than the variation in donation amounts that a donor is likely to consider: you can double your impact by doubling the amount that you give to typical charities, but you can multiply your impact by a factor of ten, 100, or even 1000 by choosing to support more effective charities [2]. In recent years, the effective altruism movement has emphasized the enormous variation in charity effectiveness and, consequently, the importance of giving more effectively [5–7] (Effective Altruism).

In addition to its potential social impact, research on the psychology of (in)effective giving is also theoretically significant. It addresses longstanding questions in moral psychology, such as: are humans truly altruistic and if so, to what extent and in what contexts [8]? What motivations and cognitive processes support altruism?

Here, we review recent research on the causes of (in)effective altruism, with a focus on charitable giving. (Other topics include volunteering and career choice.) We present a framework that distinguishes between: (i) motivational and (ii) epistemic influences on effective altruism. We also discuss (iii) interventions that could encourage more effective altruism.

Motivational Obstacles to Effectiveness

Subjectivity of Giving

Across many domains of economic life, people care about cost-effectiveness. Consumers want the most bang for their buck. Investors want higher returns. Citizens are discouraged by...
ineffective public spending. However, when it comes to charitable giving, donors seem to care less about effectiveness. Why is that?

Many people perceive charity as a matter of personal choice, such that donations need not be guided by objective measures of effectiveness [9]. For many, a donor choosing a charity is more like a gourmand choosing a restaurant than a doctor prescribing a treatment. When presented with charities addressing different causes, people tend to choose the charities that have the most personal appeal, even when informed that other options are more effective [9]. Moreover, third parties approve of donation decisions based on emotional connection more than they approve of donation decisions based on objective measures of effectiveness [9]. The idea that charity choice is a matter of subjective preference is widespread, but some philosophers have challenged this idea, arguing that if one is going to help others, one is obliged to do so as effectively as possible (Box 2). While addressing this philosophical controversy is beyond the scope of this review, as a psychological matter the perceived subjectivity of giving may be the most fundamental obstacle to effective giving.

If choosing among charities is perceived as merely an expression of personal preference, donors may see little reason to counteract the misconceptions and cognitive biases that impede effective giving. Likewise, there is little need for them to conduct research or seek expert advice to guide their decisions. In a survey of 3000 American donors [10], only 33% said that they researched the charities they considered and only 9% reported giving based on evidence of relative effectiveness. This contrasts starkly with people’s behavior as consumers and investors, where people attend closely to product reviews and seek out financial advice at great expense [11].

People may view donation decisions as subjective because they do not feel responsible or accountable [12] for the effects of their donations. Many people view donating to help strangers as nonobligatory (‘supererogatory’) [13], which suggests that there is no obligation to give in a particular way. However, there are some contexts in which people feel more obliged to consider effectiveness. For instance, in preliminary research we found that people feel more obliged to maximize effectiveness when they are the only possible helpers and are therefore solely responsible for the outcome [14]. However, individual donors are typically not the only ones who can help, making effectiveness seem less important.

Narrow Affective Motivation
Donors’ reliance on subjective preferences raises the question: what produces these preferences? Research indicates that charitable giving, and helping behavior more generally, is typically driven by emotional motivators such as empathy (or sympathy) for victims [15–17] or the positive satisfaction of personally provisioning a good (i.e., ‘warm glow’) [18]. However, these affective responses are not necessarily triggered by opportunities to help effectively. On the contrary, they often hinder more effective giving [19]. We refer to these as ‘narrow’ affective motivations, contrasting them with motivations that reflect a broader integration of information about available options and their likely consequences [20].

Personal Connection
People are more inclined to support causes to which they feel personally connected or that they find personally meaningful [9]. For example, donors in wealthy nations are more likely to be personally affected by cancer than malaria, leading to greater support for charities focusing on this more personally relevant disease [21]. Moreover, they are unlikely to change their minds if informed that charity experts consider malaria charities more effective at saving lives. Likewise,
people are often drawn toward charities that help victims suffering from the same diseases or misfortunes as their friends or loved ones [22]. Similarly, people tend to prefer animal charities that help familiar companion animals, such as dogs, over charities that help farm animals, such as pigs, although charities that help farm animals tend to reduce animal suffering more effectively (Animal Charity Evaluatorsii). Finally, people are typically loyal to the charities with which they already have a relationship as a donor [10], a form of status quo bias [23].

**Narrow Moral Circle**

People tend to favor proximate beneficiaries over distant ones. First, there is spatial distance, which typically coincides with social distance: people are more willing to help others the more they feel socially connected with them [24]. Donors in the developed world, for example, prefer local over foreign charities although charities working in distant poor countries tend to be more effective [13,25–27]. Second, there is biological distance: most people prefer to help humans over animals, even when controlling for perceived cognitive capacity and capacity for suffering [28,29]. The tendency to favor humans over animals is a stable trait that correlates with tendencies that are widely regarded as biases (e.g., racism, sexism) and is reflected in donation behavior (helping humans vs animals) [28]. Third, there is temporal distance: people prefer to help current generations over future generations, as reflected in the quip ‘What has posterity ever done for us?’. While strategies to safeguard future generations are by their nature uncertain, they may still be highly effective, as the inhabitants of the future may far outnumber the living. A strategy with a 1% chance of preventing human extinction could save vast numbers of lives, making it a rational bet from an impartial perspective [30]. Although the value of such investments is difficult to quantify, there is evidence that people’s intuitive thinking about extinction events is irrationally myopic, focused on the immediate consequences of catastrophic events rather than their long-term consequences [31].

**Scope Neglect**

Few would say that the 100th life one can save is worth less than the first life one can save, yet people’s altruistic motivations do not scale proportionately with the number of beneficiaries [32]. Rather, the perceived value of each additional life saved is reduced once large numbers have been saved. This has been called ‘psychophysical numbing’ [33–35]: compassion fades with increasing numbers of victims. In one study, people were willing to donate about the same amount to help either 2000, 20 000, or 200 000 victims [36,37]. This neglect of scope may follow from our tendency to represent values using domain-general mechanisms that evolved to represent the value of goods such as food, which (unlike human lives) exhibits steeply declining marginal value [38].

Scope neglect is accentuated for single identifiable victims, which often draw more support than large numbers of ‘statistical’ victims [39–43]. This is illustrated by the famous case of ‘Baby Jessica’ who, while trapped in a well in Texas, received over US$700 000 in donations [44]. Furthermore, adding statistical information about victims to profiles of identifiable victims can even reduce people’s willingness to donate [45].

**Prioritization Aversion**

Donating effectively requires, above all else, the prioritization of charities focused on high-impact causes [5]. The most effective charities tend to work on causes that are relatively: (i) large in scope (more individuals are affected); (ii) neglected (do not receive much support), and (iii) tractable (progress can be made more easily). Charities addressing large, neglected problems for which there are (or may soon be) scalable solutions can have an impact that is orders of magnitude larger than that of typical charities.
Prioritizing some causes over others implies prioritizing some people over others – not because they are inherently less worthy, but as a form of philanthropic ‘triage’. People are averse to making such trade-offs, especially when they involve human lives [46]. Consequently, people consider it unfair to de-prioritize certain victims because others can be helped more effectively [47,48]. This is one reason why many donors prefer to split their donations across charities [49], even when some recipient charities are much less effective than others [21,50]. However, this can reduce the overall impact of one’s giving.

While people resist the deprioritization of whole causes based on effectiveness considerations, they do consider it important to avoid charities that they perceive as wasteful. That is, they deprioritize charities that use their available resources in a suboptimal way to serve their cause. For example, people consider it wrong to support a charity that relies on a relatively ineffective medication if another charity addresses the same cause with a more effective intervention [14]. Many donors are especially concerned about supporting organizations with excessively high overhead costs [51,52] (see Overhead Myth later). These tendencies indicate that donors do care about effectiveness. When asked directly, donors often identify effectiveness as a primary motivation for giving [10,21,53]. However, donors’ concern for effectiveness is rather narrow, focused on avoiding wastefulness in the service of a given cause rather than on using effectiveness as a criterion for choosing between causes [9].

Character and Reputational Benefit
One powerful driver of prosocial behavior is reputational benefit [54–56]. Giving to charity can signal good character to potential cooperation partners, but the effectiveness of one’s donations may not influence the strength of that signal, as social rewarders pay little attention to effectiveness [57]. This may be because effectiveness has historically been difficult to track, which puts a reputational premium on prosocial actions that are well-defined and highly observable [57]. This favors visible personal sacrifice over social benefit [58] and donations based on mutually salient emotional factors [59] rather than complex calculations [60]. Consistent with this, people whose donations are based on deliberation rather than empathy are viewed less positively [61]. Under prevailing norms, donors have relatively little reputational incentive to give effectively.

Epistemic Obstacles to Effectiveness
In addition to motivational factors, epistemic factors prevent people from giving effectively, even when they might be motivated to do so.

Overhead Myth
People tend to avoid charities with high overhead (administrative) costs [26,62]. This is partly because they incorrectly believe that overhead ratio tracks cost-effectiveness [62], such that money is always better spent directly on beneficiaries. While charities can overspend on overhead, research indicates that higher overhead costs are not associated with lower cost-effectiveness [62–64]. On the contrary, to maximize effectiveness charities require competent staff, good infrastructure, and self-evaluation, all of which increase overhead costs. Donors’ focus on low overhead creates perverse incentives for charities, leading them to minimize their administrative costs at the expense of effectiveness [63]. Research shows that the ‘overhead myth’ can be partly dispelled. When people are informed that overhead ratio and cost-effectiveness are not the same, many donate to more effective charities, ignoring overhead costs [21,62]. However, the effects of such information may be limited because overhead aversion has a motivational component as well. Spending on overhead can reduce the warm glow of giving [18,52] as the relationship between donor and recipients becomes
more indirect. A more general strategy for charities to mitigate the effects of overhead aversion is to have major philanthropists cover overhead costs, allowing ordinary donors to cover direct costs [52].

Quantifiability Skepticism

Many people doubt that charity effectiveness can be quantified and compared. While people readily accept that one can compare the effectiveness of charities that work on the same problem (e.g., malaria) [21], they have doubts about comparing the effectiveness of charities that work on different problems. There are, however, tools developed by economists that can be used to quantify and compare charity effectiveness. For instance, health economists conduct cost-effectiveness analysis using quality-adjusted life years (QALYs) [65] to quantify the effectiveness of alternative health interventions (Box 1) [66]. While practically challenging, cost-effectiveness analysis allows comparisons of charities addressing different causes (e.g., malaria prevention versus education). More research is needed to investigate people’s beliefs about the quantification and comparability of charity effectiveness.

Innumeracy

Maximizing the effectiveness of one’s charitable giving may require complex quantitative reasoning, or at least an understanding of certain basic statistical and economic concepts such as probability and expected value. Many people lack such knowledge or fail to transfer it from other domains (e.g., investing) [67]. For example, many donors incorrectly believe that a charity that saves one life with certainty is more effective than a charity that has a 10% chance of saving 100 lives, although the latter can be expected to save ten times more lives over time [21]. Another important but widely misunderstood concept is marginal value, considering the good that an additional donation can do [26]. When one charity is substantially more effective than the alternative charities,
one maximizes impact by donating exclusively to the most effective charity (assuming there is no substantial decline in marginal value) [68]. Many people incorrectly believe that it is more effective to split their donations across charities that vary in effectiveness [21,26,69].

People are also misled by ‘pseudoinefficacy’: if a vast number of people are threatened, people often view saving a small number of people as just a ‘drop in the bucket’ and therefore not valuable [70]. In other words, donors attend to the proportion of victims they can save rather than the absolute number [33]. This can discourage donors from supporting solutions that, while highly effective, address problems of broad scope.

Underestimation of Effectiveness Variance

Most people underestimate the variation in charity effectiveness. Experts believe that the most effective charities working on global poverty are about 100 times more cost-effective than typical charities in that domain [2]. Laypeople, by contrast, estimate that the most effective charities are only about 1.5 times more effective than average [2]. People assume that the distribution of charities by effectiveness is approximately normal (like the distribution of adult human heights) when it more closely follows a power law distribution, such that a small number of charities are vastly more effective than typical charities, much as the largest urban areas are vastly larger than typical urban areas (Figure 1). One explanation for these vast differences in effectiveness is that the charity market is, from an effectiveness perspective, highly inefficient. Donors do not penalize charities that do less good than others in the way consumers penalize companies that offer higher prices or lower quality products. As a result, many relatively ineffective charities remain in operation without improving.

Ignorance About the Most Effective Charities

Effective altruist researchers such as those at GiveWell aim to identify the world’s most effective charities. They have converged on a relatively small number of stand-out charities that focus on supporting the world’s poorest people (GiveWell’s focus), improving conditions at factory farms, and safeguarding humanity’s long-term future (e.g., from global pandemic risks). Most
people are unaware of these charities [21], which partly explains why few people support them. At the same time, people tend to overestimate the effectiveness of charities that they find personally appealing [21], especially ones that are local [27,71]. Many people also assume that charities responding to disasters are especially effective, although charities addressing ongoing problems in poor countries (e.g., neglected tropical diseases) tend to be more effective [21,72]. This may be driven by people’s greater emotional responses to victims who have experienced distinct losses compared with victims of chronic misfortune [73].

**Strategies to Increase Effective Giving**

Previous research has focused on strategies to increase people’s willingness to give to charity and the amounts that people give [4], but there is relatively little research on strategies encouraging people to give more effectively. Whether or not it is good to give more effectively, or to give at all, is a moral question [5–7] (Box 2) that we do not attempt to answer here. Nevertheless, we proceed on the assumption that saving and improving more lives rather than fewer lives is a good thing, and discuss psychologically informed strategies to encourage more effective giving.

**Information**

A natural strategy is to inform donors about which charities are most effective (as GiveWell does) and dispel misconceptions about effectiveness (e.g., the Overhead Myth Campaign). The evidence for the effectiveness of effectiveness information is mixed. Some studies have found no effect [74,75], while some of our own studies have found a large effect, increasing the proportion of effective donors from zero to 17% or even higher [21,76]. Other studies point to heterogeneity among donors, with minimal effects overall but larger effects for donors who are more altruistically motivated [77] and more educated, especially when encouraged to think more deliberately [78]. Providing more tangible details about a charity’s intervention strategies may also make giving more effective [75,79]. All studies to date use limited amounts of effectiveness information, but some donors, including donors willing to give large amounts, may be influenced by more extensive information.

**Choice Architectures and Incentives**

**Defaults**

As with other behaviors [80], setting donating as the default behavior can increase its prevalence [81]. Default settings could also be used to increase the effectiveness of giving. While default choices are by no means specific to effective giving, highly effective defaults are especially relevant because they can be justified in terms of their greater social benefits, as in the classic case of opt-out organ donation [82].

**Incentivizing**

Another general strategy to promote a behavior is to provide incentives. Consistent with this, charities benefit from donor tax deductions and often use matching campaigns [83]. Although incentivization is a general strategy, it may be especially relevant to effective giving if there is a subset of donors who are highly motivated to incentivize others to give effectively and who are therefore willing to provide matching funds for this purpose (see later regarding Giving Multiplier). Likewise, governments could provide larger incentives for effective giving through targeted tax deductions.

**Unit Asking**

Scope neglect (see earlier) is reduced when people evaluate donation options jointly rather than separately [41]. The ‘unit asking’ technique builds on this idea [84]. Potential donors are
first asked how much they would give to help one person in need before being asked how much they would give to help a group of a specified size. Unit asking tends to increase donors’ inclinations to raise their donations proportionally with group size and thus makes giving more effective.

**Splitting**

As noted earlier, people tend to split donations across multiple charities when given the option [49,85]. While the preference for splitting can reduce the effectiveness of giving (see earlier), it can also promote more effective giving if the tendency to split causes people to allot some, rather than none, of their money to a highly effective charity. As discussed earlier, people tend to support charities that are emotionally appealing, paying little attention to effectiveness. However, there is evidence that many people do care about effectiveness and that information about effectiveness can make giving more effective [2,21]. Combining these insights suggests a new strategy to increase the effectiveness of charitable giving: many donors may be amenable to splitting their donations between an emotionally appealing charity and a highly effective charity, especially if provided with effectiveness information. We have applied this technique in the design of an online donation platform (Giving Multiplier®), which uses matching funds (see earlier) to encourage donors to split their donations between a highly effective charity and their favorite charity [76].

**Philosophical Reasoning**

Philosophical arguments of the kind pioneered by Peter Singer [86] (Box 2) can increase people’s willingness to donate to an effective charity [87,88]. Likewise, veil-of-ignorance reasoning of the kind pioneered by Rawls [89] and Harsanyi [90] – asking people to consider what they would want if they did not know who they are going to be – can counteract parochial tendencies (“narrow moral circle”) and thus make giving more effective [25]. The effective altruism movement was inspired by philosophers [1,86,91,92] and has directed billions of dollars toward effective charities [93], indicating that philosophical arguments can have a powerful impact (see also Open Philanthropy vii). However, it may be that relatively few people are receptive to such arguments [87,88]. Understanding how philosophical arguments (Box 2) might influence the behavior of more typical donors is an important topic for future research.

**Box 2. Is Effectiveness Morally Obligatory?**

In one of the most influential philosophical arguments of the 20th century, Peter Singer asked readers to imagine encountering a child drowning in a shallow pond [86]. Is it morally acceptable to allow the child to drown because one does not want to ruin one’s clothes? Singer reasonably assumed that it would be wrong to allow the child to drown. However, he argued further that affluent people have a comparable moral obligation to save the lives of distant children, at least when relatively little personal sacrifice is required. While many have found this argument convincing, many more have found it unsettling: it implies that even moderately affluent people ought to give far more to charity than most of them currently give.

More recently, philosophers have considered a related moral challenge that applies even to those who reject Singer’s conclusion: if one is going to donate, does one have an obligation to choose one of the most effective charities? Parfit considered a case in which a man is about to lose both of his arms [107]. Someone can save either one arm or both arms at great cost to themselves, but the cost is the same whether they save one arm or both. Parfit grants that one is not obliged to save a stranger’s limb(s) at great cost to oneself. However, he claims that, if one decides to help, it would be perverse not to save both arms.

Applying this logic to charitable giving, Pummer argues that even if it is not obligatory to donate to charity, it is obligatory to donate effectively once one has decided to give [108]. MacAskill [109] considers an interpersonal version of Parfit’s case. Suppose that a firefighter who is charging into a burning building has time to reach only one room. One room contains one child, whereas the other contains ten. The conclusion, says MacAskill, is the same: if one is going to help, it is wrong to help in a way that helps less effectively.

In the firefighter case, the more effective option is ten times more effective. Ord’s analysis of global health interventions indicates that this ratio undersells the moral imperative to give effectively [1]. He finds that the most effective interventions can be hundreds or even thousands of times more effective than other interventions (Figure 1).
Norm Changes
One of the most robust findings in research on social influence is the power of norms [94]. As explained earlier, prevailing altruistic norms emphasize self-sacrifice, but not effectiveness. There is also a norm against publicizing one’s altruistic behavior [95,96] and (outside certain religions) there is little expectation that people donate a significant percentage of their disposable income. In the effective altruism community, these norms have shifted [5,6]. Many effective altruists publicly commit to significant life-long giving to effective charities in the hope of encouraging others to do the same (see Giving What We Can [viii], The Life You Can Save [ix]). There is also a visible commitment to evidence-based decision making and open-mindedness to new causes. It is unknown how widely such effectiveness-oriented norms could spread, but history tells us that radical norm change is possible, as demonstrated by widespread changes in moral views about slavery and racial discrimination, women’s rights, gay rights, etc. [97,98].

Concluding Remarks
Ordinary people have the power to save and transform people’s lives through effective giving, and we are beginning to understand the factors that encourage and discourage such choices, with many open questions for future research (see Outstanding Questions).

Some obstacles to effective giving are motivational: People believe that whether and how to give is primarily a matter of personal preference. Altruism tends to be motivated by narrow affective responses that are highly sensitive to personal relevance and insensitive to the scope of the problems that may be addressed. People are also averse to prioritizing some causes over others and receive little reputational benefit (and may even pay reputational costs) for prioritizing effectiveness.

Other obstacles are epistemic: people mistakenly believe that helping indirectly through organizational overhead is wasteful and that effectiveness cannot be well quantified. People misunderstand probability, fail to think in terms of marginal value, and dismiss important benefits as a drop in the bucket. Furthermore, people are unaware that charities vary enormously in their effectiveness, and do not know which charities are most effective.

There are, however, promising strategies to address these obstacles to effective giving. More accurate information can remove epistemic obstacles. Well-designed choice architectures and incentives may overcome motivational limitations, boosting people’s existing (but rarely dominant) concerns for effectiveness. Finally, in the long run, shifting norms may lead to widespread commitments to doing as much good as possible with our resources.

Acknowledgments
We thank Samantha Kassirer, Joshua Lewis, John Halstead and Matt Coleman for their valuable inputs on the manuscript. This research was funded by EA Funds (L.C.) and EA Grants (S.S.).

Declaration of Interests
The authors have no interests to declare.

Resources
http://effectivealtruism.org
https://animalcharityevaluators.org/
https://80000hours.org/
http://givewell.com/
http://overheadmyth.com/
References


61. Montealegre, A. et al. (2020) Does maximizing good make people look bad? PsyArXiv Published online April 12, 2020. https://doi.org/10.21203/rs.3uir.6021v1

