



# Livelihoods-focused Non-profit Evaluator

AUTHOR M. LAUGHLIN

REVIEW M. FAIRLESS

FEBRUARY 2025

RECOMMENDED

**Title:** Establishing a livelihoods non-profit evaluator

**Date of Publication:** February 2025

**Contributions:** The primary author for this report was Mitchell Laughlin (Ambitious Impact Research Program Fellow 2024), supported by Morgan Fairless (Ambitious Impact Research Director). Earlier scoping research was carried out by Sam Hilton.

The authors thank all stakeholders who contributed their thoughts and feedback to this investigation, including those who responded to our call for opinions.

**Ambitious Impact (AIM)** exists to increase the number and quality of effective charities improving human and animal well-being worldwide. We strive to achieve this goal through our rigorous research process and our Incubator Program, which connects talented individuals with high-impact ideas.

Our [Charity Entrepreneurship Incubation Program](#) provides potential entrepreneurs with two months of cost-covered, intensive training designed by founders for founders, along with ongoing support. Our researchers identify evidence-based, high-impact interventions and guide founders in finding co-founders to launch and scale these ideas.

**Note to readers:** Our research is designed for AIM decision-makers and Charity Entrepreneurship Incubation Program participants. It aims to identify the best ideas for our programs. Consequently, reports on those not recommended for incubation can often be less polished.

**For questions** about the content of this research, please contact Morgan Fairless at [morgan@charityentrepreneurship.com](mailto:morgan@charityentrepreneurship.com).

# Establishing a livelihoods non-profit evaluator /

## Summary

### Description

Incubating a new evaluator which would approach evaluation from a livelihoods perspective ('improving lives, not averting deaths', i.e., focusing on incomes, education, etc.). The evaluator is expected to be rigorous and produce public recommendations, like [GiveWell](#).

### Counterfactual impact

**Cost-effectiveness analysis:** We conducted a fairly conservative back-of-the-envelope calculation (56% chance of success; 70% discount on the average money moved per year by other evaluators in their first five years; average donation around 5x GiveDirectly). Under those assumptions, we expect the evaluator to be above our bar for cost-effectiveness (56 income doublings per 1,000 USD—well above our bar; see [discussion of cost-effectiveness](#)).

**Scale this charity could reach:** We are highly uncertain where a reasonable estimate for a peak in funds moved for a successful evaluator could be. Nevertheless, we expect any ceiling to be well in excess of an organization's operating expenses, and as such, do not expect this to reduce an evaluator's cost-effectiveness below our target bar.

### Potential for success

**Robustness of evidence:** Non-profit evaluators can be categorized as "giving multipliers," organizations that aim to increase the amount of money going to non-profits. To that end, we note several anecdotal cases of evaluators who direct between 6 and 31 times their operating budget to non-profit organizations. We think their success is a marginal positive update that other actors can also achieve. Still, we suspect that early movers in the space benefitted from the relative dearth of competition (see [discussion of evidence](#)).

We also note with caution that the literature examining the psychology of effective giving is mixed and tends to find that information on effectiveness does not motivate donors to give more. We think this should continue to be evaluated and used to tailor strategies for converting donors (see discussion of this [here](#)).

**Theory of Change:** The theory of change for this intervention is relatively straightforward and reflects the approach of established players like GiveWell, Giving Green, and Animal Charity Evaluators. The organization will produce and market rigorous evaluations designed to appeal to a wide audience. These credible and persuasive analyses are expected to enhance donors' understanding of what constitutes a high-impact non-profit and increase their motivation to direct funds towards effective organizations. Ultimately, this should result in more funding for impactful charities and improved well-being for their beneficiaries (see [discussion of theories of change](#)).

The greatest uncertainties in this ToC relate to the evaluator's ability to gain recognition and transform its research into donations. While existing organizations have demonstrated success in these areas, smaller actors often struggle to establish visibility and communicate the value of their work. Additionally, producing rigorous evaluations requires a high level of research expertise, as evidenced by the challenges and errors documented by evaluators in the past.

## Other

**Expert views:** We carried out several informal consultations with individuals who have expertise in the field of non-profit evaluation and affiliated efforts, including a survey consultation ( $n = 11$ ). Livelihoods was marginally the most favored option among the three we considered. Yet, there was considerable disagreement about potential funding sources, talent availability, and other implementation factors (see [discussion of expert views](#)).

**Implementation factors:** We are most concerned about talent and tractability—we think that to achieve early success, an evaluator will need an early-stage team with strong research skills and resilience. We view non-profit evaluation as a highly

technical skill, and expect that marketing those evaluations will also take considerable effort to get right (see [further discussion in the implementation section](#)).

# Establishing a livelihoods non-profit evaluator /

## Crucial considerations

### Is there a gap in the market?

We think that *many* gaps in the market could draw attention from prospective funders: there are only a handful of evaluators, and arguably only one or two very large and popular ones. Increasing the number of evaluators in the field will likely improve viewpoint diversity and plurality of approaches (see [more on the gap here](#)).

Livelihoods emerged as a strong option among the handful we considered. We think existing evaluators in this space give much greater moral weight to health-focused interventions—especially life-saving ones—over income-focused interventions. Many experts were excited by the idea of increased evaluation in this space. However, some felt the space was already well covered by existing organizations.

### Will there be sufficient donor interest?

We are unsure but marginally confident that there will be sufficient early interest in the outputs of an evaluator's work. We have already received some informal interest from two or three foundations and/or individuals who have suggested they would be interested in either funding or consuming the research produced to inform their donations. It should be noted that given more time, we would have scoped the donor market more for this type of moral view and interest. However, we note the existence of several potentially aligned funders, such as the [Livelihood Impact Fund](#), the [Agency Fund](#) (and potentially those listed on their site as [partners](#)), and other large philanthropic donors.

### Will the new organization be successful in gaining an audience?

By audience, we mean attention from the broader public. We think gaining traction within aligned audiences, such as effective altruism-aligned individuals and organizations, would account for earlier and more tractable success for the

organization. However, we have a strong expectation that the organization seeks to broaden the types of funders and individuals it appeals to. GiveWell and Giving Green have succeeded in becoming “household brands” in the charity evaluation sector which gives us some confidence. However, other evaluators have remained relatively unknown and have adopted approaches to reaching out to high-net-worth individuals or donors.

# Table of contents

<b>1</b>	<b>Background</b>	<b>6</b>
1.1	Introduction	6
1.2	Current evaluator ecosystem	7
1.3	Gaps in the evaluator ecosystem	8
<b>2</b>	<b>Evaluator options</b>	<b>9</b>
2.1	Summary	9
2.2	What was considered	9
2.3	Livelihoods Intervention Evaluator (LIE)	10
<b>3</b>	<b>Theories of change and evidence</b>	<b>12</b>
3.1	Assumptions and key factors supporting the ToC	13
<b>4</b>	<b>Expert views</b>	<b>18</b>
4.1	How we sought views	18
4.2	Expert views on scoped options	18
<b>5</b>	<b>Cost-effectiveness analysis</b>	<b>21</b>
5.1	Impact	21
5.2	Costs	22
5.3	Results	23
<b>6</b>	<b>Implementation</b>	<b>24</b>
6.1	What does working on this idea look like?	24
6.2	Key factors	25
6.3	Remaining uncertainties	31
<b>7</b>	<b>Conclusion</b>	<b>32</b>
<b>8</b>	<b>References</b>	<b>33</b>
<b>9</b>	<b>Appendix</b>	<b>39</b>
9.1	Giving What We Can's evaluator database	39
9.2	Other options considered	41

# 1 Background

**Ambitious Impact (AIM) exists to increase the number and quality of effective non-profits working to improve human and animal well-being.** AIM connects talented individuals with high-impact ideas. We give potential entrepreneurs intensive training and ongoing support to launch ideas and bring them to scale. Our research team focuses on finding impactful opportunities.<sup>1</sup>

**As part of our 2024 research agenda, we explored interventions aimed at accelerating economic growth or significantly increasing individual incomes.** In that context, we examined the feasibility of incubating a non-profit evaluator to assess charities delivering interventions aimed at meaningfully increasing the incomes of its beneficiaries. This report provides an overview of our findings.

## 1.1 Introduction

**Impact-focused non-profit evaluation involves investigating charities to assess their potential for funding.** These evaluations often include cost-effectiveness analyses, alongside reviews of the strength of the non-profit's team, the quality of evidence supporting its programs, the clarity and feasibility of its theory of change (ToC), and its capacity to effectively utilize additional funding (e.g., [GiveWell, 2024](#)).

**Evaluators play a key role in the non-profit funding ecosystem by identifying high-impact opportunities, publicly sharing findings, and enabling a more efficient market for funding to emerge.**<sup>2</sup> By publishing evaluations, they allow donors to make more informed decisions, enabling resources to be allocated effectively. Many evaluators also promote more efficient giving by endorsing certain high-impact opportunities. Successful evaluators can raise substantial

---

<sup>1</sup> To read more about our approach to selecting intervention ideas for our program, please see Murár ([2025](#))

<sup>2</sup> Impact-focused non-profit evaluators are organizations that publish evaluations of charities based on the expected cost-effectiveness of additional funding directed towards them. Other non-profit evaluators (e.g., GuideStar) tend to focus more on charities' expenses and staff compensation.

funds, redirect donations from less effective to more impactful causes, and inspire additional contributions by demonstrating potential impact.

**Given the transformative potential of evaluators, AIM has considered incubating more organizations in this space.** This report reflects a rapid and focused research process evaluating three main options within the global health and well-being area, ultimately recommending a livelihoods-focused evaluator.<sup>3</sup>

## 1.2 Current evaluator ecosystem

**Existing impact-focused evaluators cover many problem areas.** For example, GiveWell focuses on global health and well-being, Animal Charity Evaluators (ACE) evaluates the animal welfare space, and Giving Green evaluates the climate change space<sup>4</sup>.

Since their inception, the above evaluators have shifted considerable funds to their recommended charities, well in excess of their operating budgets.

**Table 1: Lifetime expenses and money moved**

Evaluator	Estimated expense spending (2023 prices)*	Estimated money moved (2023 prices)*
GiveWell <sup>†</sup>	USD 75.3 million	USD 2,175.0 million
Animal Charity Evaluators <sup>^</sup>	USD 8.8 million	USD 75.1 million
Giving Green <sup>~</sup>	USD 1.5 million	USD 19.1 million

<sup>†</sup>2007-2022. <sup>^</sup>2013-2023. <sup>~</sup>2020-2023. \* Figures adjusted by the US GDP deflator.

Other organizations exist within this ecosystem, as highlighted in [Giving What We Can's evaluator database](#), each with different worldviews, strategies, and research approaches. Additionally, many individuals (see [here](#)) and organizations, such as Rethink Priorities, have published evaluations of their own.

<sup>3</sup> Many more options could exist for consideration, and in fact we may return to this question in a more systematic fashion in future. Options that were less considered but could have a positive outlook include preventing cases of extreme suffering, systemic change, agency-improving interventions, democracy, and evaluating effective giving organizations.

<sup>4</sup> We focused on these three examples given their similarities to our vision of what an AIM-incubated evaluator would like. That is, their main remit is producing public evaluations focused on the most cost-effective charities in a particular cause area.

## 1.3 Gaps in the evaluator ecosystem

**More non-profit evaluators could address several gaps in the existing evaluator ecosystem:**

**Limited capacity:** Existing evaluators cannot assess all high-impact charities due to resource constraints, potentially leaving impactful organizations underfunded.

**Lack of second opinions:** Many charities are evaluated by only one organization, giving donors few alternative perspectives to compare.

**Inconsistent evaluation methods:** Where multiple evaluations do exist, different approaches and worldviews can make it difficult for donors to draw clear comparisons.

**Underrepresentation of income-based interventions.** Many evaluators prioritize health-focused interventions over income-based ones, neglecting donors with alternative priorities.

**Insufficient evaluations for emerging opportunities ([Giving What We Can, 2024](#)):**

As funding in the effective giving<sup>5</sup> ecosystem grows, the marginal cost-effectiveness of established high-impact charities may decline, increasing the need for evaluations to identify emerging or overlooked opportunities.

---

<sup>5</sup> 'Effective giving' is the practice of donating to charities in a way that uses evidence and reason to maximize, or meaningfully increase, the positive impact of each dollar given.

## 2 Evaluator options

### 2.1 Summary

**After scoping options and incorporating experts' feedback, we explored the merits of incubating a livelihoods interventions evaluator (LIE).** This organization would evaluate the impact of charities with interventions aimed at increasing the quality of life of individuals by (primarily) raising the incomes of beneficiaries—sometimes referred to as focusing on “improving lives rather than averting deaths.” It would seek to capture donors interested in contributing to global well-being from an income-focused, rather than health-focused, perspective, noting the relative neglectedness of these interventions amongst existing evaluators like GiveWell and Founders Pledge.

### 2.2 What was considered

**Throughout the research process, we shallowly considered multiple evaluator options to recommend for incubation.** We assessed options on their suitability for incubation in AIM's [Charity Entrepreneurship Incubation Program](#) and their perceived likelihood of success.

**Early in the research process, we decided to constrain the scope of our research and consider only evaluators within the global health and well-being space.** This ruled out evaluators that could focus on, for example, artificial intelligence safety opportunities or effective-giving organizations.

Three options were scoped for feedback from the staff of existing evaluators and grantmakers, as well as the wider Ambitious Impact community.<sup>6</sup> These were:

1. **Livelihoods intervention evaluator (LIE):**<sup>7</sup> evaluating charities delivering interventions aiming to increase the incomes and quality of life of beneficiaries

---

<sup>6</sup> The scoping document distributed for feedback is available [here](#).

<sup>7</sup> In the scoping document we presented to experts this was originally referred to as an income growth intervention evaluator.

2. **Macroeconomic growth intervention evaluator (MGE):** evaluating funding opportunities to accelerate economic growth in low- and middle-income countries
3. **Early-stage and local non-profit evaluator (ELCE):** evaluating early-stage global health and well-being-focused charities, including local organizations that may not be well known to international funders

**After incorporating feedback, we decided to prioritize focusing on the livelihoods intervention evaluator (LIE).** This feedback is contained in [Section 4](#).

## 2.3 Livelihoods Intervention Evaluator (LIE)

**A livelihoods intervention evaluator (LIE) would assess the impact of charities implementing interventions aimed at significantly increasing the incomes of beneficiaries.**

Table 2 below provides examples of problem areas and charities that could fall within the LIE's scope.

**Table 2: Potential focus areas for LIE**

Problem area / Sector	Existing organization
Agriculture	One Acre Fund
Microfinance	Small Enterprise Foundation
Entrepreneurship	Village Enterprise
Poverty graduation	BRAC
Small and medium enterprise (SME) support	Acumen Fund
Distribution of specific products	Deworm the World
Unconditional cash transfers (UCTs)	GiveDirectly
Education	TaRL Africa
Other	Malengo

### 2.3.1 Donor market

A LIE would seek to attract donors interested in contributing to global development from an income-focused rather than health-focused perspective.

**Existing evaluators in this space, including AIM's research team to some extent, place significantly greater moral weight on health-focused interventions—particularly life-saving ones—compared to income-focused interventions.**

GiveWell cost-effectiveness models estimate that averting one death of a child under five is approximately 117 times more valuable than doubling consumption for one person for one year ([GiveWell, 2020](#)). Similarly, Founders Pledge views averting an under-five death as 129 times more valuable than doubling one person's income for a year ([Founders Pledge, 2022](#)).

**Donors who prioritize increasing income (or consumption) over health outcomes may be encouraged to donate more—and, in their worldview, more effectively—when more evaluations of income-focused charities are available.**

This perspective aligns with the funding priorities of organizations like [The Livelihood Impact Fund](#), [Mulago Foundation](#)'s education, livelihood, and energy programs, and [The Life You Can Save](#)'s education and quality of life funds. This highlights the potential appeal of a LIE to donors who share these values.

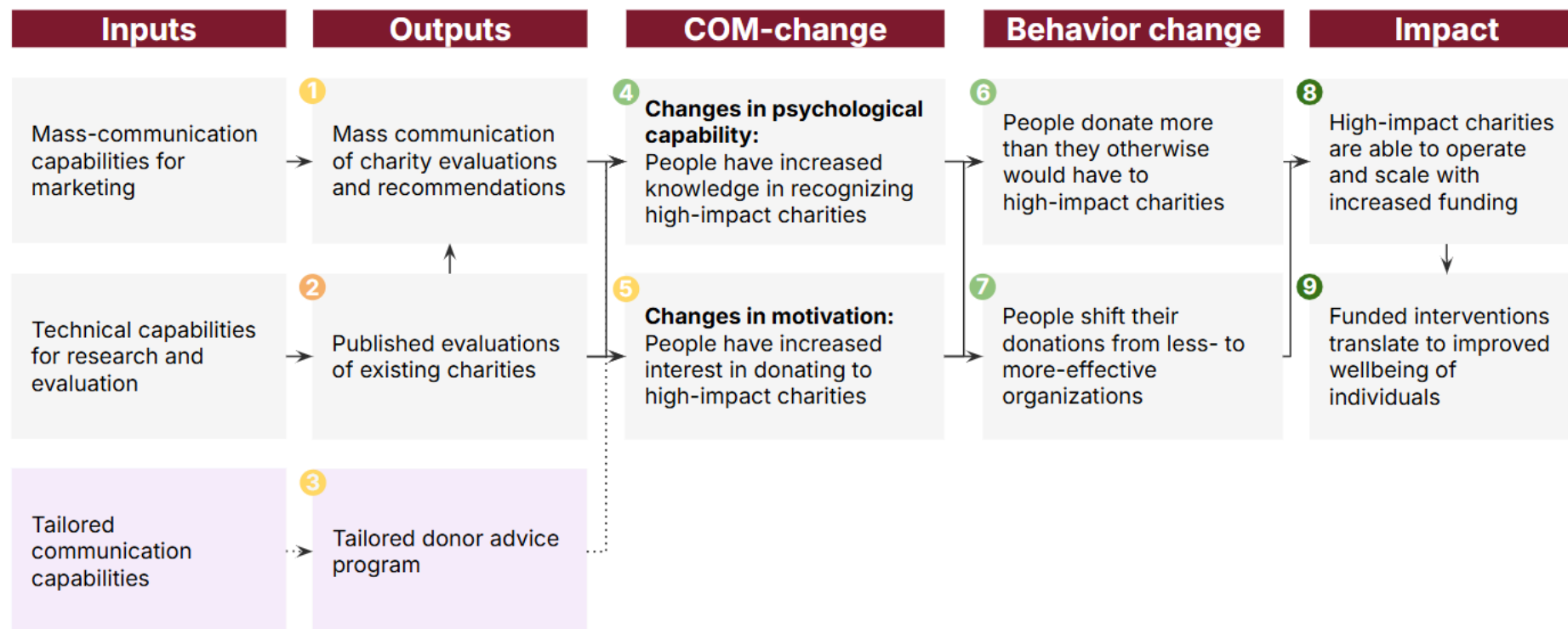
### 2.3.2 Neglectedness

Evaluators like GiveWell, Founders Pledge, SoGive, The Life You Can Save, and the Happier Lives Institute have evaluated charities addressing the problem areas listed in Table 1. However, none of these organizations has made this their primary focus, and, to our knowledge, no organization currently produces evaluations primarily of charities delivering income growth-focused interventions.

Many experts were excited by the idea of increased evaluation in this space. However, some experts felt these organizations already covered the space well.

### 3 Theories of change and evidence

**This section provides an overview of the ToC for this idea.** It represents a broad depiction of our thinking behind how this intervention works. The diagram below attempts to capture the high-level ToC of this non-profit.<sup>8</sup>



**Figure 1: The primary ToC of this non-profit.** Note that those steps in mauve are considered optional in this ToC.

Scale: key uncertainty, high uncertainty, some uncertainty, low uncertainty, unconcerning.

<sup>8</sup> COM refers to Capability, Opportunity and Motivation

### 3.1 Assumptions and key factors supporting the ToC

1. The evaluator can successfully mass-communicate themselves and their research.
2. High-impact organizations exist and the evaluator can conduct evaluations.
3. The evaluator can deliver a tailored donor advice program for high-net-worth individuals.
4. Evaluator's publications and media give its audience a better understanding of constitutes a high-impact non-profit.
5. Evaluators' publications and media create more motivation in their audience to donate to high-impact charities.
6. The mentioned changes in psychological capability and motivation are sufficient to encourage non-donors to start donating and existing donors to increase their contributions.
7. The mentioned changes in psychological capability and motivation are sufficient to prompt individuals to reallocate their donations toward more effective interventions.
8. Funding is a current constraint to many current and prospective high-impact organizations.
9. The funded interventions translate to improved well-being of individuals.

Scale: *key uncertainty*, *high uncertainty*, *some uncertainty*, *low uncertainty*, *unconcerning*.

#### 1. The evaluator can successfully market their research

We have some concerns about whether the evaluator would be able to market their work successfully. This would likely be dependent on the communication or

marketing skills of the co-founding team or an early hire, and the quality of the published evaluations.

As we articulated in our report on exploratory altruism ([Hilton & Agarwalla, 2021](#)), “there is a strong precedent for new non-profit evaluators in new causes to attract significant attention.” This precedent is evident in several notable examples:

- “Animal Charity Evaluators preceded and helped seed the effective animal advocacy movement” ([Hilton & Agarwalla, 2021](#)).
- Giving Green launched successfully, referenced in articles published in [The Atlantic](#), [Forbes](#), and [The New York Times](#), and two ([here](#) and [here](#)) in [Vox](#) within its first year ([Lawsen, 2021](#)).
- GiveWell has an extensive list of links to articles written about its work on its “[What Others Are Saying](#)” page, with four from its first year.

However, there is likely significant survivorship bias in assessing the above examples, and we think that some non-cause-area-specific evaluators, for example, have had less success in marketing their research and building an audience.<sup>9</sup>

## 2. The evaluator can conduct evaluations

**We have moderate concerns that a founding team would be able to produce high-quality evaluations.** While many organizations currently produce, in our opinion, quality evaluations, nearly all of them have received criticism for their research output in the past. Some examples of criticism are expanded in more detail in [Section 6.2.1](#) and [Section 6.2.6](#). These organizations have strengthened their research methodologies in response to this criticism, though it is likely that these research errors have hampered their growth and impact.

Ultimately, we believe that the quality of research will be heavily dependent on the skillset of the founding team and early hires of an evaluator.

---

<sup>9</sup> “Survivorship bias is a cognitive shortcut that occurs when a successful subgroup is mistaken as the entire group, due to the invisibility of the failure subgroup” ([Pilat & Krastev, 2020](#)).

### **3. The evaluator can deliver a tailored donor advice program for high-net-worth individuals**

**At least in the short term (within five years), we are uncertain whether a founding team could successfully balance delivering tailored donor support alongside its broader remit, raising some concerns with this assumption.**

For example, in its “Change of Direction” update GiveWell noted that its approach to raising funds through high-net-worth donors was too time-intensive and that it should instead focus on broadening and deepening its research ([GiveWell, 2008](#)). Our understanding is that GiveWell did not prioritize donor outreach until 2016, nine years after its founding ([GiveWell, 2016](#)). Around 12 percent of GiveWell’s named staff now work in its Philanthropy team ([GiveWell, 2024](#)).

Many organizations, including some effective giving organizations incubated by AIM, successfully deliver tailored donor advice to high-net-worth individuals. As such, we believe this is achievable, but may be better prioritized after the early years of an evaluator’s life.

### **4. Publishing and marketing high-quality evaluations leads to increases in the knowledge of donors**

**We assume that individuals engaging with an evaluator's publications will better understand non-profit cost-effectiveness.** Multiple papers have shown that donors respond positively to increased information on non-profit cost-effectiveness, and this assumption seems intuitively straightforward.

### **5. Publishing and marketing high-quality evaluations increases donors’ motivation to support high-impact charities**

**We have some concerns about whether marketing an evaluator’s research will increase donor motivation to donate to highly effective non-profits.** Aside from [Karlan and Wood \(2014\)](#) ( $n = 21,643$ )—which suggests that providing evidence of a non-profit’s cost-effectiveness may reduce the motivations of certain “warm glow”

donors<sup>10</sup>—there is little evidence on how information about a non-profit's cost-effectiveness impacts donor motivation to give effectively.

**Some evidence supports that engaging with arguments similar to those of philosopher Peter Singer can increase donations to effective charities ([Lindauer et al., 2023](#),  $n = 1,260$ ; [Schwitgebel, 2020](#),  $n \approx 2,500$ ).** An evaluator could reference such arguments in publications but would need to do so cautiously to avoid deterring potential donors.

## 6. Changes in motivation and knowledge lead to donors donating more than otherwise

We assume that an evaluator's work would lead to an aggregate increase in donations.

Non-profit evaluators have a strong record of acting as "*giving multipliers*: funding opportunities which, as their main output, cause more money to be given to high-impact interventions and organizations" ([Flynn-Do, 2022](#)). In 2022 (its most recent year for which records are available), GiveWell directed USD 439 million, 31 times its 2022 expenses ([GiveWell, 2023](#)). Giving Green's 2023 multiplier for money influenced on operating expenses was 19 ([Giving Green, 2024](#)), and we estimate ACE's was around 6 ([ACE, 2024](#)).<sup>11</sup> We believe that much of this funding has been raised counterfactually—that is, it would not have been raised without their work.

While we are uncertain as to whether an additional incubator would have as much success—given that we believe existing evaluators filled comparatively more neglected niches—we still expect a livelihoods evaluator to move significant funding to impactful non-profits.

There is little formal literature exploring the success of high-impact non-profit evaluators, though there is some literature exploring the psychology of effective giving. Laboratory experiments by [Metzger and Günther \(2019a\)](#) ( $n = 2,523$ ) and

---

<sup>10</sup> In this paper, those donors are less driven by altruism and more by the positive feeling of having donated, regardless of the expected impact of the donation.

<sup>11</sup> ACE switched to reporting financial year figures from 2022-23. For comparability we estimated calendar year figures. These are visible in our [CEA](#).

[Metzger and Günther \(2019b\)](#) ( $n = 265$ ) found that information about an organization's effectiveness did not affect average donation levels to the organization. Similarly, [Karlan & Wood \(2014\)](#) ( $n = 21,643$ ) found that impact information did not affect average donation amounts, though with considerable heterogeneity across donors, with large prior donors likely to give more and small prior donors giving less often. This provides some evidence against our assumption, though we perceive the anecdotal successes of prior evaluators to be worth greater weight.

## **6. Changes in motivation and knowledge lead to donors shifting their donations from less effective to more effective non-profits**

**We assume that published evaluations would lead to donors shifting their funding from less to more effective donation opportunities.** Numerous individuals and organizations already fund the proposed problem areas with the intent to maximize impact or support highly impactful non-profits. Many evaluators have successfully redirected funding to cost-effective non-profits. Therefore, we assume that these donors would be receptive to reallocating their funding toward more cost-effective opportunities if made aware of them.

**In a literature review of the psychology of effective giving, [Caviola et al. \(2021\)](#) report that the “evidence for the effectiveness of effectiveness information [on increasing effective giving is mixed.” (p. 602)** As with assumption 5., this challenges our prior beliefs. However, we consider the past experiences of evaluators to be more relevant to our ToC than the findings of this literature.

## 4 Expert views

### 4.1 How we sought views

Before narrowing our research scope to focus solely on the global health and well-being space, we engaged in conversations with experts about the possibility of founding an evaluator across various problem areas.

After defining this focus, we sought feedback on the three ideas outlined in [Section 2](#) from staff at existing evaluators, grantmakers, and the broader AIM community.

### 4.2 Expert views on scoped options

**Expert opinions varied significantly**, with some options receiving strong support from certain individuals while garnering minimal support from others.

**However, on balance, experts expressed optimism about incubating a non-profit evaluator, with income growth interventions emerging as the most promising focus area.**

#### 4.2.1 Income growth intervention evaluator (IGE)

**Among the options considered, experts were most optimistic about AIM incubating an IGE.**

**Supporters noted that interventions improving livelihoods are relatively neglected by current evaluators and believed donors would value this focus.** Others argued that organizations like GiveWell and Founders Pledge already covered the space well.

**One expert expressed the need for a founding team with strong credentials and research skills**, while others emphasized the importance of effective branding for

an evaluator of this type. Additionally, one expert suggested framing the organization as a “livelihoods” intervention evaluator.

#### 4.2.2 Macroeconomic Growth Intervention Evaluator (MGE)

Many experts expressed uncertainty around incubating an evaluator focused on macroeconomic growth.

A common view was that researching this problem area would be particularly challenging due to a weaker evidence base and poor feedback loops. Additionally, some suggested that, given Open Philanthropy (OP) is launching a similarly focused program ([O'Keeffe-O'Donovan et al., 2024](#)), it would be best to wait and consult with OP on their experiences before incubating in this area. We agree with this assessment.

Additionally, one expert raised that Ambitious Impact may have reputational risks if such an organization was perceived as attempting to influence LMIC policy.

One expert also suggested that the income and macroeconomic growth evaluators could be merged. This idea has not been explored in this report, but an IGE founder could consider the merits of this approach after assessing OP's and Founders Pledge's experiences.

#### 4.2.3 Early-stage and local non-profit evaluator (ELCE)

**Experts held mixed views on an ELCE and shared different visions for how one might operate.** This may have been a fault of reasoning transparency in our scoping document.

Experts noted the potential actual or perceived conflict of interest associated with incubating an organization that would likely evaluate many other AIM-incubated charities. Also, many questioned the merits of establishing an evaluator to publish assessments on small charities and instead suggested establishing public funds or working with existing organizations (e.g., D-Prize) and funds to more directly improve the cost-effectiveness of grants. One expert shared this sentiment who

expressed it would be difficult to build a reputation and network around an early-stage evaluator with a broad remit. Many experts also expressed concerns that a local non-profit evaluator would struggle with language barriers.

Two experts also questioned whether "early-stage" and "local" paired well together. One of these experts suggested that there could be merits in establishing a national non-profit evaluator. An evaluator of this type would aim to identify cost-effective charities operating within a specific country, catering to donors who were interested only in funding non-profits within their borders and attempting to build a greater culture of focus on cost-effectiveness.

#### **4.2.4 General views**

Some experts' feedback shared a theme of whether establishing a public fund would better serve the aims of a proposed evaluator. Some experts also questioned why existing evaluators and research organizations could not expand their program areas to cover the perceived gaps in the evaluation space. Experts also widely expressed that AIM must prioritize recruiting individuals with strong research skills, which may complicate the standard Charity Entrepreneurship Incubation Program recruitment process.

## 5 Cost-effectiveness analysis

**As with many of our meta-non-profit models, we would characterize this modeling as closer to a back-of-the-envelope approach, given its speculative nature.** We have grounded our estimates in available data, the model is relatively simplistic and highly sensitive to our subjective assumptions.

Notably, we measured outcomes across only the first five years and estimated the likelihood of success (which we usually don't do for our direct-delivery models). We felt there was little benefit to extending our model past this point for several reasons:

- We assumed that an unsuccessful evaluator would shut down before its fifth year, and so extending the model would only favor the case for incubating an evaluator.
- In our base case estimates the evaluator was many times as cost-effective as GiveDirectly in expectation—well above AIM's target bar—and extending the model would only strengthen this case.
- We are more uncertain of estimates in later years.

We also present results of the same model, assuming a 100% chance of success and extending the analysis to all the years for which precedent data is available.

Our cost-effectiveness model is available [here](#).

### 5.1 Impact

The modeled impacts rely on the following inputs:

1. **Total money moved.** Our estimates for total money moved in each year took the arithmetic mean of public money moved estimates from GiveWell, Animal Charity Evaluators, and Giving Green in their first years of existence.
2. **Replicability adjustment factor.** We adjusted our estimates for 'Total money moved' to account for expectations that a new evaluator would likely not be

as successful as GiveWell, Giving Green, and ACE. AIM occasionally utilizes an adjustment factor of 69% from [Coville and Vivalt \(2017\)](#), and we anchored to this factor for our CEA. While we appreciate this factor is not intended for our purposes in this case, as it is based on papers in development economics, we feel our assumption is reasonable.

3. **New money raised as a percent of total money moved.** It was subjectively assumed that 35% of the total money raised by Effective Giving Initiatives (EGIs) represented “new money”—funds that would not have been raised without their efforts. This is a reduced figure compared to a previous AIM estimate of 50% ([Baas et al., 2023](#)). The remaining 65% of the “total money moved” was “shifted money”—funds that would have been raised regardless of EGI’s involvement. Changes in this proportion did not significantly affect the model’s sensitivity.
4. **Income doublings per dollar of charity funds reallocated.** This was estimated as the income doublings per dollar of GiveDirectly, based on GiveWell’s November 2024 estimates (see [here](#)).
5. **Income doublings per dollar of evaluator’s recommended charities.** We assume that recommended charities will be around 10 times as cost-effective as GiveWell’s pre-November 2024 estimate for GiveDirectly’s cost-effectiveness. This assumption was informed by Founders Pledge’s evaluation of Imagine Worldwide—a Malawian education non-profit—which estimated it as 10.9 times as cost-effective as GiveDirectly ([Albinsky, 2023](#)).

Impact was estimated in terms of income doublings. The annual impact was the sum of the income doublings bought with “new money raised” and income doublings gained from shifting money from less effective to more effective interventions.

## 5.2 Costs

Costs were modeled as the estimated operating expenses of GiveWell across its first five years of operations. These were then multiplied by a factor of 1.5 to adjust

for our view that an organization would need to hire experienced and credentialed researchers, which would likely increase staff costs.

## 5.3 Results

**Our model projected a cost-effectiveness of 56 income doublings per USD 1,000 in operating expenses. This equates to a cost-effectiveness of USD 41 per DALY equivalent.** These estimates incorporate tractability assumptions. These are visible in our model.

**The results of our cost-effectiveness analysis should be treated with caution, given the subjectivity of our assumptions.**

**Table 3: Alternative models**

	Income Doublings per USD 1,000	USD per DALY-equivalent
5-year data only		
Assuming success	70	33
Limited tractability model (~56%)	56	56
16-year data		
Assuming success	213	11
Limited tractability model (~56%)	211	11

## 6 Implementation

### 6.1 What does working on this idea look like?

The work of a non-profit evaluator will be split between producing and marketing research.

**The first year of an evaluator will likely involve the following:**

- **Scoping:** Identifying a sector (see Table 2) to focus on for the first stage of non-profit evaluations (our soft sense is that education could be a good space to start in given donor interest) → Identifying the most promising programs within that area → Identifying charities for evaluation that deliver these programs.
- **Evaluating:** Assessing charities' impact by analyzing their evidence of effectiveness, cost-effectiveness, and capacity to utilize additional funding. Clear and effective communication with evaluated charities will be crucial to minimize grievances.
- **Recommending:** Publishing the findings of the evaluations with recommendations for donors interested in funding work in this area. Likely promoting some organizations as, e.g., "top charities" or "our recommendations".
- **Impact tracking:** Measuring money moved and other success indicators. This complex task can be supported through affiliate links, direct donations to recommended charities via an on-site donor platform, feedback from charities and donors, media mentions of the evaluator and its research, and metrics like audience size at speaking appearances (e.g., podcasts and conferences).

**Once an evaluator is especially comfortable with the quality of its initial recommendations it could increase its marketing efforts.** This could entail, e.g., writing opinion or guest pieces for existing media, appearing in interviews or on

podcasts, hosting Q&As of research and/or organizational philosophy, presenting at conferences or attending other in-person events, or developing a tailored donor advice program. However, we would caution that some of our team believe past evaluators have initially over prioritized promotion to the detriment of research quality.

**Subsequent evaluation rounds could investigate other sectors. Some donors are interested in funding sector-specific opportunities**, with donors of this type spread across multiple sectors. As such, an evaluator may move more in funding by having recommendations covering at least some of these different problem areas.

**It should be noted that this strategy is not clear-cut.** GiveWell initially followed this approach, recommending its recommended “top charities” and certain less-exceptional, but still highly cost-effective by typical standards, “standout charities”. They later discontinued recommending standout charities: “We think our standout charities are excellent, but we believe donors should support top charities” ([GiveWell, 2022](#)). Giving Green similarly concluded that it over-recommended offset donation options, ultimately detracting from total donor impact ([Giving Green, 2022](#)).

## 6.2 Key factors

This section summarizes our concerns (or lack thereof) about different aspects of a new non-profit putting this idea into practice.

**Table 4: Implementation concerns**

Factor	How concerning is this?
Talent	Moderate concern
Access to information	Low-moderate concern
Access to relevant stakeholders	Low concern
Feedback loops	Low concern

Factor	How concerning is this?
Funding	Low-moderate concern
Neglectedness	Low-moderate concern
Tractability	Moderate concern
Complexity of scaling	Low concern
Risk of harm	Low concern

## 6.2.1 Talent

**This non-profit will require certain co-founder skills.** Primarily, it would be beneficial for the founding team to have strong research skills and a background in producing evaluations of charities or interventions. The research skillset of at least one co-founder (or an early hire) would ideally be comparable to that of a senior researcher at GiveWell.

**In addition to research experience, we expect that strong marketing and communications skills would be beneficial for developing and growing an audience.** Strong interpersonal skills would also be beneficial in delivering tailored advice to high-net-worth donors and in communicating with research partners and stakeholders from evaluated charities.

**Given the research skills required to succeed in this area, we have moderate concerns about this non-profit's talent requirements.**

**Table 5: Founder requirements and nice to haves**

Must have	Preferable (offsets a 10% diff in incubatee strength)	Preferable, all else equal
From <a href="#">GiveWell's Senior Researcher job description</a> : <i>"quantitatively-oriented advanced degrees or an</i>	<ul style="list-style-type: none"> <li>Development economics background (either in academia or an organization of</li> </ul>	<ul style="list-style-type: none"> <li>Communications or marketing background.</li> <li>Strong interpersonal skills.</li> </ul>

Must have	Preferable (offsets a 10% diff in incubatee strength)	Preferable, all else equal
<i>undergraduate degree and substantial, relevant experience using empirical tools to make rigorous, evidence-based decisions in the real world".</i>	comparable reputation to the World Bank). <ul style="list-style-type: none"> <li>• Skills in non-profit evaluation.</li> </ul>	

## 6.2.2 Access

### Information

**We believe the information required to conduct quality non-profit evaluations would mostly be available** from charities and the development economics literature.

### Relevant stakeholders

**We expect that this non-profit will operate collaboratively with other relevant stakeholders.** As such, it will be crucial to be able to access and build positive rapport with them. Stakeholders could include other research organizations (e.g., J-PAL and GiveWell), grantmakers (e.g., The Livelihood Impact Fund), and the charities the organization evaluates.

We do not expect that developing rapport with these stakeholders will be particularly difficult for a founding team.

**However, maintaining strong relationships with charities will be vital.** In their "Our Mistakes", GiveWell and Giving Green list shortfalls in managing relationships with evaluated charities<sup>12</sup>. An evaluator should aim to communicate clearly with its evaluated charities and balance transparency with respecting the wishes of the organizations it assesses. Failing to maintain relationships with charities would

<sup>12</sup> See GivingGreen's "Publication of "shallow dive" write-ups of organizations that are not part of our top recommendations" ([here](#)) and GiveWell's "Not informing candidate charities of our recommendation structure prior to publishing recommendations" ([here](#)).

likely hamper an evaluator's ability to access information relevant to its research. On the other hand, overly compromising transparency would likely harm the evaluator's reputation, and could worsen the quality of its recommendations and donors' decisions.

### 6.2.3 Feedback loops

We are not overly concerned about feedback loops for this non-profit. We expect the non-profit directors will know within twelve months of publishing their first set of recommendations whether their evaluations are influencing funding decisions.

#### Tracking impact

Money moved would be a key metric, and could be tracked through affiliate links, by talking to the charities and funders, or by moving money through the organization. Readership of ideas and other outreach measures (such as press mentions) could also be tracked.

Animal Charity Evaluators' [Giving Metrics Report](#) details its approach to estimating and tracking impact, including metrics such as money moved, donor growth and retention, and web traffic. A founding team could consult other evaluators like ACE to learn effective strategies for early-stage impact tracking.

### 6.2.4 Funding

**We have some concerns that an organization would struggle to secure funding in the medium term (i.e., in its third and later years).**

We expect funding to be available for the first one to two years of an evaluator's life. If the evaluator has not demonstrated impact by then (e.g., because it had prioritized research over marketing—which may be a viable strategy) it may struggle to attract further funding.

The impact-focused philanthropic space has relatively high salaries for researchers, so we anticipate that this organization would have greater operating expenses than a standard AIM-incubated meta-non-profit given stricter talent

requirements. Having said that, we do not think that funding requirements will exceed a few million dollars a year at maximum unless the non-profit does very well and needs to scale massively.

In the short term, we expect there to be funding available for work of this type from within the AIM Network, such as the seed funding circle. A few donors have expressed interest in supporting this type of initiative.

### 6.2.5 Neglectedness

**We have some concerns as to whether this space is neglected.** While we ultimately believe it is currently underexplored by evaluators and that an evaluator in this space could be highly impactful, expert opinion was divided. Some experts felt that GiveWell and Founders Pledge already cover this space sufficiently, suggesting further evaluation might not increase funding to high-impact non-profits. Others, however, argued this space remains neglected, and would benefit from greater evaluation.

### 6.2.6 Tractability

**We have moderate concerns regarding the tractability of an evaluator.** Our concern pertains largely to the ability of an evaluator to conduct quality evaluations.

**Producing non-profit evaluations is not trivial, and most impact-focused evaluators have been critiqued for their research quality in the past.** Some cases are listed below, though it should be noted that examples were readily available largely because of these organizations' transparency and their openness to feedback.

The Happier Lives Institute (HLI) has thrice revised its cost-effectiveness estimates for its evaluation of the organization StrongMinds ([Happier Lives Institute, 2023](#)). Following criticism ([Snowden, 2023](#); [GiveWell, 2023](#)) of its evaluation suggesting StrongMinds was roughly 9.5 times as cost-effective as GiveDirectly, HLI has downgraded its estimate for StrongMinds to 3.8 times as cost-effective as Give Directly. HLI has bolstered its research team and methodology in response, and acknowledges evaluation errors in its "[Our Blunders](#)".

Giving Green (GG) was critiqued in early 2021 for its general research methodology and specific recommendations ([Lawsen, 2021](#)). For its initial recommendation of the organization Sunrise, GG also acknowledges this as a mistake in its "[Our Mistakes](#)".

Animal Charity Evaluators (ACE) has also been critiqued for their research ([Nathan, 2016](#); [Halstead, 2018](#); [Anon., 2022](#)) and their recommendations were not relied upon by Giving What We Can following its 2023 evaluations of evaluators ([Giving What We Can, 2023](#)).

GiveWell lists many mistakes in its "[Our Mistakes](#)", many relating to research errors. GiveWell also held a [Change Our Mind Contest](#), inspired by a critique from HLI ([McGuire et al., 2022](#)), with HLI's critique and the contest's submissions leading to changes in GW's cost-effectiveness estimates across multiple evaluations.

This concern links heavily to our concerns regarding talent (as in [Section 6.2.1](#)).

## 6.2.7 Complexity of scaling

**We are unconcerned about the complexities associated with scaling.** We do not expect an evaluator to be limited by funding or evaluation opportunities. In a 2023 blog post, GiveWell articulated that it expected its research to continue to yield more opportunities than it could fund ([Arjmand, 2023](#)).

However, we have some concerns as to how much funding an evaluator could move at its peak. GiveWell's year-on-year growth in funds moved has slowed considerably across its life. Similarly, Animal Charity Evaluators has seen some stagnation in growth in its money influenced across the past five years.

We are highly uncertain as to where a reasonable estimate for a peak in funds moved for a successful evaluator could be. Nevertheless, we expect any ceiling to be well in excess of an organization's operating expenses, and as such do not expect this to reduce an evaluator's cost-effectiveness below our target bar.

### 6.2.9 Risk of harm

**We consider the risk of harm to be low.** While we originally had concerns that an evaluator might redirect funding from more cost-effective health-focused organizations to less cost-effective livelihood-focused organizations., most experts viewed this risk as minimal. Consequently, we have reduced our level of concern.

## 6.3 Remaining uncertainties

Given more time, we would have liked to have further explored how much demand there was for an evaluator in the space of income-focused interventions. Specifically, we would ideally have a better understanding of how much funding is currently directed toward charities seeking to increase the incomes of their beneficiaries and, on average, how effective that funding is.

It would have also been useful to discuss directly with evaluator founders what they felt were the key challenges to launching a successful evaluator.

## 7 Conclusion

After reviewing this research, our decision board recommended this idea for incubation in the CEIP.

We think a pair of well-resourced co-founders with acumen for research and marketing should launch a non-profit evaluator that focuses on interventions that try to improve the livelihoods of people, in particular those living in low-resource contexts. While there are a handful of strong evaluators in the field, their moral weights often (not always) value saving lives more over improving them, which may not cater to the moral commitments or areas of focus of philanthropic funders. Therefore, we are excited by the prospect of contributing to a more vibrant and diverse market of evaluators that, as a whole, contributes to increased donations to the very best non-profit work out there.

## 8 References

- 2020 update on GiveWell's moral weights. (2021, November). GiveWell.  
<https://docs.google.com/document/d/1hOQf6Ug1WpoicMyFDGoqH7tmf3Njjc15Z1DGERaTbnI/edit?tab=t.0>
- 2023 GiveWell cost-effectiveness analysis—version 4 (public). (2023). GiveWell.  
[https://docs.google.com/spreadsheets/d/18ROI6dRdKsNfXg5glyBa1\\_7eYOjowfbw5n65zkrLnvC/edit?gid=1680005064#gid=1680005064](https://docs.google.com/spreadsheets/d/18ROI6dRdKsNfXg5glyBa1_7eYOjowfbw5n65zkrLnvC/edit?gid=1680005064#gid=1680005064)
- Albinsky, V. (2023). Are education interventions as cost effective as the top health interventions? Five separate lines of evidence for the income effects of better education [Founders Pledge]. *Effective Altruism Forum*.  
[https://forum.effectivealtruism.org/posts/8qXrou57tMGz8cWCL/are-education-interventions-as-cost-effective-as-the-top#VII\\_Discussion](https://forum.effectivealtruism.org/posts/8qXrou57tMGz8cWCL/are-education-interventions-as-cost-effective-as-the-top#VII_Discussion)
- Annual Reports. (2024). Animal Charity Evaluators.  
<https://animalcharityevaluators.org/about/impact/annual-reports/>
- Anonymous. (2022, September 2). An Evaluation of Animal Charity Evaluators. *Effective Altruism Forum*.  
<https://forum.effectivealtruism.org/posts/pfSiMpkmskRB4WxYW/an-evaluation-of-animal-charity-evaluators>
- Arjmand, I. (2022, September 6). Announcing the Change Our Mind Contest for critiques of our cost-effectiveness analyses. *The GiveWell Blog*.  
<https://blog.givewell.org/2022/09/06/announcing-change-our-mind-contest/>
- Arjmand, I. (2023, April 10). How much funding does GiveWell expect to raise through 2025? *The GiveWell Blog*.  
<https://blog.givewell.org/2023/04/10/expected-funds-raised-through-2025/>
- Baas, J., Speziali, F., & Moore, L. (2023). *Effective Giving Incubation Report*. Giving What We Can & non-profit Entrepreneurship.  
[https://9475dbf4-555e-4808-9886-5f8ee815cc82.usrfiles.com/ugd/9475db\\_d06fa170d29349318fe61cd87f1d5faa.pdf](https://9475dbf4-555e-4808-9886-5f8ee815cc82.usrfiles.com/ugd/9475db_d06fa170d29349318fe61cd87f1d5faa.pdf)
- Barnes, T. (2024, August 1). *Lead Exposure Elimination Project*. Founders Pledge.  
<https://www.founderspledge.com/research/lead-exposure-elimination-project-leep>
- Berman, J. Z., Barasch, A., Levine, E. E., & Small, D. A. (2018). Impediments to Effective Altruism: The Role of Subjective Preferences in Charitable Giving. *Psychological Science*, 29(5), 834–844. <https://doi.org/10.1177/0956797617747648>

- Bettle, R. (2022, June 13). *Suvita*. Founders Pledge.  
<https://www.founderspledge.com/research/suvita>
- Bettle, R. (2023, June). *Replicability & Generalisability: A Guide to CEA discounts*. Founders Pledge.  
[https://docs.google.com/document/d/1eJBsmNG-iRJ-twoHaoztQUmB4pmEHut8\\_oT3xgtldnl/edit?tab=t.0#heading=h.7om527x8q9t7](https://docs.google.com/document/d/1eJBsmNG-iRJ-twoHaoztQUmB4pmEHut8_oT3xgtldnl/edit?tab=t.0#heading=h.7om527x8q9t7)
- Bettle, R., Albinsky, V., & Barnes, T. (2022, August). *Moral Weights*. Founders Pledge.  
<https://docs.google.com/document/d/1VZ8VGly3bZELUZ6Xwz7uPBEawvzFYX83m5Vn7iUHN-s/edit?tab=t.0#heading=h.y8bv2big6pot>
- Caviola, L., Schubert, S., & Greene, J. D. (2021). The Psychology of (In)Effective Altruism. *Trends in Cognitive Sciences*, 25(7), 596–607.  
<https://doi.org/10.1016/j.tics.2021.03.015>
- Caviola, L., Schubert, S., & Nemirow, J. (2020). The many obstacles to effective giving. *Judgment and Decision Making*, 15(2), 159–172.  
<https://doi.org/10.1017/s1930297500007312>
- Caviola, L., Schubert, S., Teperman, E., Moss, D., Greenberg, S., & Faber, N. S. (2020). Donors vastly underestimate differences in charities' effectiveness. *Judgment and Decision Making*, 15(4), 509–516.  
<https://doi.org/10.1017/s1930297500007452>
- Changelog*. (2023, November 27). Happier Lives Institute.  
<https://www.happierlivesinstitute.org/changelog/>
- non-profit evaluation*. (n.d.). Effective Altruism Forum. Retrieved November 13, 2024, from <https://forum.effectivealtruism.org/topics/non-profit-evaluation>
- Cohen, A. (2023, March 21). *Assessment of Happier Lives Institute's Cost-Effectiveness Analysis of StrongMinds*. GiveWell.  
<https://www.givewell.org/international/technical/programs/strongminds-happier-lives-institute>
- Family Empowerment Media*. (2021, December 21). Founders Pledge.  
<https://www.founderspledge.com/research/family-empowerment-media>
- Financials*. (2024). Animal Charity Evaluators.  
<https://animalcharityevaluators.org/transparency/financials/>
- Flynn-Do, K. (2022, August). *Cause Area Summary: Giving Multipliers*. Founders Pledge.

<https://docs.google.com/document/d/1ZoODVspd7x36uxIxNG6Ke1-6IZXOLnuL0A8VDWFIHlw/edit?tab=t.0>

*GiveWell, CEA of GiveDirectly's Cash for Poverty Relief program, 2024 (public).* (2024, November). GiveWell.

<https://docs.google.com/spreadsheets/d/1LqkkwWCud9qGjSJmw9re6DIm5ocpPOaqOJ-KIEVLppk/edit?usp=sharing>

*GiveWell Incubation Grants.* (2023, May). GiveWell.

<https://www.givewell.org/research/incubation-grants>

*GiveWell Metrics Report - 2022 Annual Review.* (2023). GiveWell.

[https://files.givewell.org/files/GiveWell\\_Metrics\\_Report\\_2022.pdf](https://files.givewell.org/files/GiveWell_Metrics_Report_2022.pdf)

*GiveWell's Impact.* (2023, December). GiveWell.

<https://www.givewell.org/about/impact>

*Giving Green's 2023 annual impact report.* (2024, March 14). Giving Green.

<https://www.givinggreen.earth/post/2023-annual-report>

*Giving Metrics Report.* (2022). Animal Charity Evaluators.

<https://animalcharityevaluators.org/about/impact/giving-metrics/>

Gosnell, G., & Basnak, M. (2023, May 16). *Family Empowerment Media: track record, cost-effectiveness, and main uncertainties.* Rethink Priorities.

<https://rethinkpriorities.org/research-area/family-empowerment-media/>

*Gross Domestic Product: Implicit Price Deflator.* (2024, October 30). Federal Reserve Bank of St. Louis. <https://fred.stlouisfed.org/series/GDPDEF>

Halstead, J. (2018, September 18). Concerns with ACE research. *Effective Altruism Forum.*

<https://forum.effectivealtruism.org/posts/zri7LsivEAqnYZ6Wt/concerns-with-ace-research>

Hassenfeld, E. (2022, October 28). We're discontinuing the standout non-profit designation. *The GiveWell Blog.*

<https://blog.givewell.org/2021/10/05/discontinuing-standout-non-profit-designation/>

Hilton, S., & Agarwalla, V. (2021). *Exploratory Altruism.* non-profit Entrepreneurship.

[https://3394c0c6-1f1a-4f86-a2db-df07ca1e24b2.filesusr.com/ugd/9475db\\_a215cf49667a40f7aca07db9c10bd246.pdf](https://3394c0c6-1f1a-4f86-a2db-df07ca1e24b2.filesusr.com/ugd/9475db_a215cf49667a40f7aca07db9c10bd246.pdf)

- Hoeijmakers, S., Townsend, M., Horowitz Friedman, A., & Giving What We Can. (2023, November 23). GWWC's evaluations of evaluators. *Effective Altruism Forum*. [https://forum.effectivealtruism.org/posts/PTHskHoNpcRDZtJoh/gwwc-s-evaluations-of-evaluators#Animal\\_non-profit\\_Evaluators\\_\\_ACE\\_](https://forum.effectivealtruism.org/posts/PTHskHoNpcRDZtJoh/gwwc-s-evaluations-of-evaluators#Animal_non-profit_Evaluators__ACE_)
- Imagine Worldwide: Public*. (n.d.). Founders Pledge. Retrieved November 13, 2024, from [https://docs.google.com/spreadsheets/d/1V-Dk9VeaRaSgbak9C-Ube1paUALtPASByZ\\_DdiMUX78/edit?gid=262547608#gid=262547608](https://docs.google.com/spreadsheets/d/1V-Dk9VeaRaSgbak9C-Ube1paUALtPASByZ_DdiMUX78/edit?gid=262547608#gid=262547608)
- Job Application for Senior Researcher at GiveWell*. (2024). GiveWell. [https://job-boards.greenhouse.io/givewell/jobs/4253692008?gh\\_src=57add528us](https://job-boards.greenhouse.io/givewell/jobs/4253692008?gh_src=57add528us)
- Karlan, D., & Wood, D. H. (2017). The effect of effectiveness: Donor response to aid effectiveness in a direct mail fundraising experiment. *Journal of Behavioral and Experimental Economics*, 66, 1–8. <https://doi.org/10.1016/j.socec.2016.05.005>
- Lawsen, A. (2021, January 21). Why I'm concerned about Giving Green. *Effective Altruism Forum*. <https://forum.effectivealtruism.org/posts/7yN7SKPpL3zN7yfcM/why-i-m-concerned-about-giving-green>
- Lindauer, M., Mayorga, M., Greene, J., Slovic, P., Västfjäll, D., & Singer, P. (2020). Comparing the effect of rational and emotional appeals on donation behavior. *Judgment and Decision Making*, 15(3), 413–420. <https://doi.org/10.1017/S1930297500007208>
- Metzger, L., & Günther, I. (2019a). Is it what you say or how you say it? The impact of aid effectiveness information and its framing on donation behavior. *Journal of Behavioral and Experimental Economics*, 83. <https://doi.org/10.1016/j.socec.2019.101461>
- Metzger, L., & Günther, I. (2019b). Making an impact? The relevance of information on aid effectiveness for charitable giving. A laboratory experiment. *Journal of Development Economics*, 136, 18–33. <https://doi.org/10.1016/j.jdeveco.2018.08.015>
- Moore, L., & Hoeijmakers, S. (2024, September 30). Updates on the effective giving ecosystem (MCF 2024 memo). *Effective Altruism Forum*. <https://forum.effectivealtruism.org/posts/eoAuq8M8fydGBd5jv/updates-on-the-effective-giving-ecosystem-mcf-2024-memo>

- MVP evaluator database.* (2024, July). Giving What We Can.  
<https://givingwhatwecan.notion.site/MVP-evaluator-database-c5a60dfe0b9d4f3c9e7cda2d54e3d318>
- Nathan, H. (2016, December 10). *The Actual Number is Almost Surely Higher.*  
<https://medium.com/@harrisonnathan/the-actual-number-is-almost-surely-higher-92c908f36517>
- O'Keeffe-O'Donovan, R., Latham-Proença, R., & Kim, O. (2024, October 16). *Economic growth in low- and middle-income countries.* Open Philanthropy.  
<https://www.openphilanthropy.org/research/economic-growth-in-low-and-middle-income-countries/>
- Our impact.* (2024). Giving Green. <https://www.givinggreen.earth/our-impact>
- Our Mistakes.* (2024, November 1). Giving Green.  
<https://www.givinggreen.earth/post/giving-green-s-mistakes#viewer-dlcle>
- Our People.* (2024). GiveWell. <https://www.givewell.org/about/people#Philanthropy>
- Our Progress to Date.* (2021, March). GiveWell.  
<https://www.givewell.org/about/progress#Changeofdirectionpublished11172008>
- Pilat, D., & Krastev, S. (2020). *Survivorship bias.* The Decision Lab.  
<https://thedecisionlab.com/biases/survivorship-bias>
- Process for Identifying Top Charities.* (2024, June). GiveWell.  
<https://www.givewell.org/how-we-work/process>
- Research Overview.* (2024, August). GiveWell. <https://www.givewell.org/research>
- Sarek, K., & non-profit Entrepreneurship. (2022, February 2). CE is scaling – our 3-year plan and current hiring round. *Effective Altruism Forum.*  
<https://forum.effectivealtruism.org/posts/EL3dKu5XcfjYhdjFX/ce-is-scaling-our-3-year-plan-and-current-hiring-round>
- Savoie, J. (2022, February 13). We need more nuance regarding funding gaps. *Effective Altruism Forum.*  
[https://forum.effectivealtruism.org/posts/GFkzLx7uKSK8zaBE3/we-need-more-nuance-regarding-funding-gaps#Organizational\\_size\\_variation](https://forum.effectivealtruism.org/posts/GFkzLx7uKSK8zaBE3/we-need-more-nuance-regarding-funding-gaps#Organizational_size_variation)
- Savoie, J. (2024, January 12). A short comparison of starting an effective giving organization vs. founding a direct delivery non-profit. *Effective Altruism Forum.*  
<https://forum.effectivealtruism.org/posts/MayBxEme6mWn9fswG/a-short-comparison-of-starting-an-effective-giving>

- Schwitzgebel, E. (2020, June 23). Contest Winner! A Philosophical Argument That Effectively Convinces Research Participants to Donate to non-profit. *The Splintered Mind*.  
<https://schwitzsplinters.blogspot.com/2020/06/contest-winner-philosophical-argument.html>
- Snowden, J. (2023, February 25). Why I don't agree with HLI's estimate of household spillovers from therapy. *Effective Altruism Forum*.  
<https://forum.effectivealtruism.org/posts/h5sJepiwGZLbK476N/assessment-of-happier-lives-institute-s-cost-effectiveness>
- Taimaka. (2024, September 26). Founders Pledge.  
<https://www.founderspledge.com/research/taimaka>
- Zarosa, U., & non-profit Entrepreneurship. (2023, March 6). After launch. How are CE charities progressing? *Effective Altruism Forum*.  
<https://forum.effectivealtruism.org/posts/cHDz2R5FfWGoZgWoZ/after-launch-how-are-ce-charities-progressing>

## 9 Appendix

### 9.1 Giving What We Can's evaluator database

**Table 6: Evaluators**

Name	Cause areas	Public non-profit recommendations?	Host funds ?	Estimated number of FTE involved in evaluations	Year of founding
GiveWell	GH&W	Yes	Yes	~30	2007
Founders Pledge	Multiple	Yes	Yes	~10	2013
Animal Charity Evaluators	AW	Yes	Yes	~5	2013
Longview Philanthropy	Longtermism	No	Yes	~5	2018
Effective Altruism Funds	Multiple	No	Yes	~5	2017
Happier Lives Institute	GH&W	Yes	No	~5	2019
Giving Green	Climate	Yes	Yes	~5	2020
The Life You Can Save	GH&W	Yes	Yes	<3	2013
doebem	GH&W	Yes	Yes	<3	2017
Center on Long-Term Risk	Longtermism	No	Yes	<3	2013
Centre for Exploratory	Multiple	Yes	No	<3	2022

Name	Cause areas	Public non-profit recommendations?	Host funds ?	Estimated number of FTE involved in evaluations	Year of founding
Altruism Research					
non-profit Box	GH&W	Yes	Yes	<3	2020
SoGive	Multiple	Yes	No	<3	2016
Focus Philanthropy	AW	No	No	<3	2023
Open Philanthropy	Multiple	No	No	~50	2014

Note: GH&W—Global Health and Well-being, AW—Animal Welfare, Longtermism—Long-term future.  
Source: Giving What We Can's [MVP evaluator database](#). Accessed 4 November 2024.

## 9.2 Other options considered

**Other options were also suggested by experts and lightly considered by AIM researchers.** These were:

- *Meta evaluator*: evaluating the giving multipliers of effective giving organizations (e.g., Giving What We Can and Doneer Effectief)
- *Mental health evaluator*: evaluating organizations and opportunities aimed at improving the mental health of beneficiaries
- *Systemic change evaluator*: evaluating organizations and opportunities aimed at driving large-scale social change
- *Artificial intelligence safety evaluator*: evaluating organizations and opportunities aimed at minimizing existential risks from AI

### **Early-stage and local non-profit evaluators (ELCE)**

An early-stage and local non-profit evaluator (ELCE) would evaluate organizations delivering interventions in global health and well-being (GH&W) with lesser track records and on a smaller scale than those that would typically be evaluated by GiveWell. For more information, see Joey Savoie's 2022 post on the necessity of nuance when discussing funding gaps, specifically his example under the 'Organizational size variation' heading ([here](#)).

There are two types of early-stage non-profits that we have in mind as potential targets of evaluation:

- Early-stage non-profits that are well-connected to Effective Altruism and thus known to evaluators (not requiring much scoping)
- Early-stage non-profits working on the ground with little connection to Effective Altruism (requiring scoping from evaluators)

### **Donor market**

An ELCE would seek to capture donors looking to support charities that have the potential to be more cost-effective than GiveWell's top charities, or donors interested in supporting early-stage charities. These donors would be expected to have a higher risk tolerance concerning their donations and would be less focused on the track records of an evaluated non-profit.

Outside of the EA movement, there is also a growing interest in supporting locally-led organizations. This evaluator could tap into this interest and help direct funding to highly effective organizations.

## **Neglectedness**

Currently, no organizations publicly rank or evaluate early-stage GH&W charities by cost-effectiveness.

However, some organizations produce occasional public evaluations of early-stage charities, and some organizations provide fellowship programs where early-stage charities can seek funding and support.

### *Effective Altruism connected non-profits*

There seem to be many organizations working in this space, though none in an especially focused manner.

- Rethink Priorities has previously completed an [evaluation](#) of Family Empowerment Media (2023), which was funded by an anonymous donor.
- Founders Pledge has evaluated and recommended many early-stage AIM-incubated charities, including:
  - Taimaka (incubated 2024 [founded 2019], evaluated [2024](#)),
  - LEEP (incubated 2020, evaluated [2023](#)),
  - FEM (incubated 2020, evaluated [2022](#)), and
  - Suvita (incubated 2019, evaluated [2022](#)).
- GiveWell provides funding to high-impact opportunities, and has provided, what were previously referred to as Incubation Grants to early-stage charities in the following categories ([GiveWell, 2023](#)):
  - Academic research to evaluate program evidence in order to build the evidence base for a potential priority program or to further our understanding of our current priority programs.
  - Early-stage funding for a promising organization.
  - Monitoring and evaluation of an existing organization.

Below is a (non-exhaustive) list of organizations working in this space:

- Charity evaluators

- GiveWell
- Founders Pledge
- Rethink Priorities

*Non-Effective Altruism connected non-profits*

Many organizations have non-profit fellowship programs, which early-stage charities can apply to, though none publish cost-effectiveness analyses of the charities in question.

- D-Prize—[D-Prize](#)
- Y Combinator's Nonprofit Track—[YC's Non-profit Program](#)
- Mulago Foundation—[Rainer Arnhold Fellows](#)
- Echoing Green—[Fellowship](#)
- The Draper Richards Kaplan Foundation (DRK)—[What We Fund](#)
- Global Innovation Fund (GIF)—[What We Fun](#)

