

CHARITY RANKINGS: DELIVERING DEVELOPMENT OR DEHUMANISING AID?

LOGAN COCHRANE^{1*} and ALEC THORNTON²

¹*University of British Columbia (Okanagan), Kelowna, BC, Canada*

²*The University of New South Wales, UNSW Canberra, Australia*

Abstract: Individuals want to know which organisations to donate to, and a variety of organisations have developed ranking systems to guide them. This paper explores charity ranking, with a particular focus on the increasing role of impact and ‘cost-effectiveness’. Ranking systems are composed of a selection of metrics, which may miss important components and, as a result, create a set of unintended outcomes. We argue that an emphasis on cost-effectiveness and impact in ranking promotes simple, technocratic activities, negatively affects human rights-based interventions and de-prioritises inventions that work in remote, complex settings. The topic of charity ranking and its influence on private donors is limited in the literature, and this paper seeks to make a contribution to this important debate. Copyright © 2015 John Wiley & Sons, Ltd.

INTRODUCTION

Relationships between charity rankings, aid selection, quality and effectiveness receive scarce attention in the development studies literature. This paper brings to light some of the issues embedded within charitable ranking and the processes by which they operate, particularly as it relates to defining ‘effectiveness’, determining which metrics are utilised and how. We refer to the individuals, groups and organisations that have sought to rank charities and charitable action as the charity rankers. These charity rankers make evaluations in order to develop rankings for the purposes of influencing donor practice. Evaluation and metrics are not neutral, objective calculations of reality, they are selective and can have unintended outcomes. For example, donor-prescribed and/or donor-required reporting metrics can shape project design, implementation and reporting (Ebrahim, 2010; Buffardi, 2011; Andrews, 2013; Mathenge, 2013). In this paper, we focus on charity rankers that seek to directly influence the decision-making of individual donors and provide a detailed case study of one charity ranker.

*Correspondence to: Logan Cochrane, University of British Columbia (Okanagan), 3333 University Way, Kelowna, BC, Canada. E-mail: logan.cochrane@gmail.com

As the charity rankers influence billions of dollars of donations annually, the choice of metrics and processes used to evaluate charitable organisations require greater attention and reflection. While there are organisations that rank and promote charitable activity without being based upon rigorous evaluation, these will not be analysed in detail, nor will official development assistance, which is well served in the development literature (Fuchs *et al.*, 2014), and is not the primary focus of charity rankers. The charity rankers referred to in this paper are those organisations and groups that seek to influence individual giving by offering research-based evaluations of charitable activity.

This paper critiques the way ‘cost-effectiveness’ has become the foundation of many organisations that rank charitable activity and charitable organisations. Our objective is not to offer authoritative answers on how charity ranking ought to or ought not to be done, it is to further a discussion about the processes and impacts of charitable ranking, identify areas that require additional reflection and provide concrete examples of the concerns being raised. For example, the exclusion of human rights from charity-ranking assessments may entrench disempowerment and further marginalisation as cost-effectiveness models promote programming that provides services to those easiest to reach or for whom change requires low-cost interventions. Rather than view all people as equal and equally deserving of receiving the benefits of charitable activity, or by prioritising the needs of the most vulnerable, the cost-effectiveness model may solidify and/or widen inequalities. In discussing the problematic components embedded within the approaches and metrics utilised in charity ranking, this paper contributes to an ongoing discussion about what ‘effectiveness’ means and how the shift towards cost-effective evaluation requires further consideration and analysis. Re-evaluation of ‘effectiveness’ is particularly important as charitable ranking, collectively, influences billions of dollars in charitable giving. Charity Navigator alone, for example, suggests that it impacts US\$10billion in charitable giving annually.

An emphasis on quantifiable indicators to measure the effectiveness of western-led development strategies emerged in the 1960s and, following criticisms from a Marxist ‘cultural turn’ (Harvey, 1982, 2003; Barnes, 2001), became reinvigorated with the rise of neoliberalism in the 1980s and 1990s (Murray & Overton, 2011). Hermes and Lensink (2001), in their critique of a World Bank report on aid assessment (Dollar & Pritchett, 1998), argued that the debate on aid had shifted, whereby ‘effectiveness’ was dependent on ‘good governance’, or that ‘effectiveness’ of a recipient policy regime was a facilitator of economic growth. Subsequently, some argue that the debate has settled on a positivist macroeconomic dogma that without aid, growth and poverty reduction may stagnate (McGillivray *et al.*, 2006; Mekasha & Tarp, 2013). The emergence of neostructuralism in the 2000s tempered the neoliberal market-led approach with a renewed emphasis on the role of the state in determining its own development, including aid delivery (Murray & Overton, 2011). Nevertheless, technocratic indicators remain common in the aid environment, resulting in donors and recipients adopting them in assessing the impact and cost-effectiveness of their aid spending.

In deconstructing a ‘cost-effective’ approach, we agree with Kothari (2005: 443) in her critique of the increasing ‘professionalisation of international development’, which promotes a ‘technical process of intervention that maintains the legitimacy and authority of Western modernity and the dominance of the neoliberal agenda’. Technocratic evaluations ignore the role of states in influencing aid flows, as well as aid spending and aid reporting by local recipients and undervalue donor decisions regarding aid selection and quality. The processes that render development technical have direct and indirect impacts on how, when, why and

for whom development occurs (Rose, 1999; Li, 2007). It is our view that the broader trends embedded within processes of neoliberalisation and professionalisation are reflected in many of the charity-ranking methodologies, which are founded in assumptions that development can be best achieved with technocratic solutions and best determined with the calculation of costs and impacts. In making this argument, we do not claim that these processes are inappropriate in all situations; rather, we draw attention to what is excluded when such processes are utilised, such as human rights, and the unintended impacts of them. In this regard, we align ourselves with Mosse (2004, 2010) and Eyben (2000, 2011) in our critique of assumptions embedded within development and the trends of the sector.

The approach that emphasises ‘cost-effectiveness’ as the primary mechanism for decision-making will be visited throughout this paper, as this is linked to positivist and neoliberal shifts in thinking about ranking interventions by evaluators (Collier & Dollar, 2001). These evaluators tend to formulate their assumptions using a limited set of criteria that aim to classify a diverse range of non-profits into pre-determined categories (Lowell *et al.*, 2005). The increasing popularity of these technocratic approaches may have unintended outcomes, which often go unexplored by both the ranking bodies and the donors. The charity-ranking process often does not consider sociocultural and political contexts, human rights, need prioritisation, ownership, sustainability and unintended outcomes (Jones, 2011).

THE CHARITY RANKERS

The charity rankers are often non-profit organisations themselves, with private funding sources. These groups create a set of metrics by which charitable activity can be evaluated and then ranked accordingly. The metrics, processes and ranking systems are diverse, with some focusing on organisational processes and others on the results of particular activities. What unites the charity rankers is that they convey their findings using the language of scientific research, suggesting the conclusions are objective observations of reality. In practice, however, the number of organisations evaluated is limited, severely restricting what is considered, and the metrics define what components of that charitable activity is valued. While hundreds of charity rankings exist, only a few have the resources available to undertake research of a sufficiently large body of material to conduct rigorous evaluations (Mitchell, 2014). Many of the smaller organisations rely upon these better-funded organisations, often repeating and slightly modifying the results of other charity rankers. Charity Navigator and GiveWell are two of the most influential and well-resourced organisations that stand out amongst the collection of charity rankers in this regard.¹ Both organisations have significantly altered individual giving, potentially influencing billions of dollars of donations annually, as well as the enterprise of charity ranking in totality, with the former promoting the assessment of administration costs and the latter introducing results-based evaluation.

Currently, Charity Navigator is the largest and most influential charity-ranking service (founded in 2001). It seeks to guide ‘intelligent giving’ by rating charity performance with

¹While conducting the initial research for this paper, we assessed 30 charity rankers based on their evaluation methodologies, metrics and the ranking conducted. That initial assessment highlighted Charity Navigator and GiveWell as the two most prominent organisations (in 2013), being widely referred to and having the resources and capacity to conduct significantly more original evaluations.

the criteria of accountability, financial health and transparency. It wants to facilitate the development of a more efficient and responsive philanthropic marketplace with its choice of metrics and processes of evaluation. In response to recent shifts in the charity-ranking sector, Charity Navigator is in the process of building in a new dimension to their analysis, focusing on results reporting. Donors, practitioners and academics have criticised Charity Navigator's ranking system for not assessing results or impact (Lowell *et al.*, 2005; Heiberg, 2009). Those challenges and critiques will not be repeated here.

The number of charity rankers is large, exceeding at least one hundred (Mitchell, 2014), each of which will not be analysed here. This paper focuses primarily upon one case study, GiveWell, selected for its unique approaches, influence in affecting private donor giving and limited analysis in the academic literature. GiveWell (founded in 2007) ranks charities by track record, cost-effectiveness, room for more funding (a measure of the effectiveness of additional donations), as well as transparency and accountability. GiveWell's influence has rapidly expanded, and its innovative metrics and evaluations have influenced the way charity ranking is conducted. Before GiveWell entered the charity-ranking sphere, administrative costs were the main metric used to convey ranking information, with Charity Navigator being the most influential of those. For the most part, the emphasis in ranking has shifted to cost-effectiveness and impact (or results). Discussed later, the results of our analysis of GiveWell will be contrasted with a third organisation, GlobalGiving, which is not presented as a charity ranker, as it does not rank charities, but as an alternative approach for engaging donors.

GlobalGiving, founded in 2002, is unique in its development of a platform for communities and organisations to propose projects and donors to fund projects under a 'cost-effective banner', such as emergency food relief or mosquito nets, while also providing a means for a wide range of other potentially high-impact projects to be proposed and funded, such as female empowerment, provision of clean water and advocacy. What makes this platform unique is that it screens locally-based and proposed projects and facilitates a system whereby donors can follow their particular regional and programming interests. In this regard, GlobalGiving is neither offering top-down technocratic suggestions nor ones that are bound by a certain set of predetermined effectiveness criteria. Because GlobalGiving does not rank charity, we analyse its platform as an approach to engaging donors, rather than as a charity ranker.

METHODOLOGY

During the last couple of decades, a wide variety of individuals and groups have emerged whose purpose is to direct private donors to charities or projects viewed as being highly cost effective. This is done by ranking charities, projects and intervention types using a range of criteria. Such efforts have greatly impacted private donor giving and the charitable sector—in both positive and negative ways. The ranking systems differ, regarding the types of metrics, evaluations and rankings they use in making recommendations. This paper will explore the seemingly neglected components of charity ranking and offer examples that demonstrate why greater attention ought to be paid to these areas. Because of the large number of individuals, groups and organisations ranking charities, this paper is qualitative in nature and utilises a detailed case study of GiveWell to explore theoretical concerns associated with technocratic charity-ranking processes. This organisation was selected for detailed analysis because of its influence, innovative practices and limited

discussion in the academic literature. GiveWell has significantly increased its ability to influence private donor giving in recent years but has not been subject to the same level of analyses in the academic literature.² Although not explored to the same extent, the platform developed by GlobalGiving is contrasted with the model of GiveWell in the concluding sections of this work.

Because of the case study approach of this paper, the conclusions ought to be understood within this context and not as generalised findings that apply to all charity rankers, their evaluations and use of metrics. This approach seeks to highlight the existence of unintended outcomes as a result of charity-ranking processes and, in particular, the negative impact these may have upon human rights. On this note, this criticism does not advocate that charity rankers are intentionally acting against the promotion and protection of human rights, rather than their evaluations and metrics have unintended impacts of this nature. The changing environment of the charity-ranking sector is noteworthy, as rankers adjust their rankings and re-evaluate on an ongoing basis. Therefore, the foci of this paper are the evaluations, metrics and processes that existed during 2013 and 2014.

DEVELOPMENT AID: THEORY AND PRACTICE

The relationships between development theory, strategy and the aid sector are complex and in flux. In 1955, Arthur Lewis suggested that the main impediment to development was a lack of capital, and Millikan and Rostow (1957) agreed, adding that technical assistance was a key supportive component. Conditional aid, best exemplified in structural adjustment programmes of the 1980s, advocated a set of requirements that nations needed to adopt in order for development to occur and for funding to be granted. Random control trial (RCT)-based development, becoming influential in the 2000s, sought to find the greatest impact interventions and utilised an evaluation approach common in clinical research. Although diverse, these examples share two common traits in their underlying assumptions about how development ought to function; first, that development is driven by the donor and, second, that a technocratic, positivist solution is able to solve the challenge at hand.

Critique of donor-driven development is not new. Rosenstein-Rodan (1943) emphasised the role of partnership and recipient-driven development, and Pearson (1969) argued that, far too often, services given by donors did not match the context within which they operated. However, critiques of technocratic assumptions that form the basis of charitable ranking have faced less scrutiny. This is partly due to scale; while Pearson and Rosenstein-Rodan focused upon official development assistance, the charity rankers assessed here seek to influence individual giving practices. Understanding the unintended outcomes of charity ranking, and what the important missing factors might include, require greater reflection, analysis and critique. Voices of concern in this regard are not absent (Deaton, 2010; Eyben, 2011), although much of the continued critique has focused on official development assistance.

²As an indicator of the level of academic research, a search on Google Scholar (17 August 2015) for 'Charity Navigator' resulted in 1380 results, while 'GiveWell' resulted in 402.

MEASURING AID EFFECTIVENESS: AN OVERVIEW

Governance of aid in the 1980s was dominated by a top-down structure promoting austerity measures or ‘structural adjustment policies’ created by international financial institutions, namely, the World Bank and the International Monetary Fund. State recipients agreed to International Monetary Fund/World Bank terms to restructure economies in order to access loans. By the early 1990s, a shift from state-centred aid to a more project-based structure developed, seeking a larger role by civil society (e.g. NGOs). However, aid policy is ultimately driven by western market ideology, the ‘Washington Consensus’, which effectively requires recipient states to provide favourable market conditions, in the form of neoliberal restructuring, as a conditionality of aid. As a result, this conditionality re-introduced the role of the state in the development agenda.

By the 2000s, the Millennium Development Goals established measurable indicators of symptoms, as opposed to the causes, of poverty that helped to provide clear objectives for development aid to potentially target (Overton *et al.*, 2013). Along with the 2005 Paris Declaration, which originally sought to empower civil society, the development agenda of the last decade has contributed to a neo-structural turn that re-empowers the state as the primary development agent; re-emerging as the primary distributor of aid, as well as a provider of economic and social development. Thus, aid is subject to ‘process conditionalities’ (Gould, 2005), involving donor management auditing and reporting for development spending that is ‘cost effective’. Alternatively, the ‘managerialism’ of aid is preferred amongst aid reformists seeking ‘aid-free’ solutions through deeper engagement with open market and free trade policies (Gulrajani, 2011). From a Foucauldian perspective, increasing ‘managerialism’ in aid reinforces a distorted hierarchy between donors and recipients that perpetuates a cycle of underdevelopment, where the latter are deemed incapable of directing their own social, political and economic development (Gulrajani, 2011).

Cost-effectiveness is generally understood to be an assessment of the costs of an intervention and the change that resulted, based upon a set of criteria or indicators. This might, for example, consider the cost of a project and the resulting increase in school enrolment. This paper does not argue that processes of evaluation of this sort are resulting in incorrect assessments. Rather, it takes a critical view of an apparent dominance of econometrics in measuring aid (Collier & Dollar, 2001; Easterly & Williamson, 2011; Knack *et al.*, 2011; McGillivray, 2003). Broad-based applications of econometrics are indicative of a positivist, ‘top-down’ neoliberal trend to quantify, generalise and thus reduce complexities inherent in aid selection, quality and delivery. In a similar fashion, emphasising cost-effectiveness appears to minimise the importance of social welfare impact at the local level for effective project implementation, which the donor funding facilitates (Lowell *et al.*, 2005; Jones, 2012). Conceptually, ‘effectiveness’ is not particularly harmful; rather, it is the processes that determine the nature of effectiveness and how certain definitions of effectiveness can negatively affect beneficiaries, projects and the sector. The applicability of econometrics is likely limited to localised, small project evaluation; therefore, its power to generate ‘universal’ lessons is unlikely.

INFLUENCING GIVING

With annual private donations in the billions (Global Humanitarian Assistance, GHA, 2012), charitable giving is big business. Donor giving spreads across all socioeconomic

sectors of society, for example, approximately 90% of Americans give to charities annually (DellaVigna *et al.*, 2012). Most charitable giving is based on personal interest, preference and/or experience with a particular organisation (Breeze, 2010). Donors are also pushed to give by emotional appeals (Singer, 2009). For example, several major natural disasters in 2010 led to an increase of 23% of public and private humanitarian giving (GHA, 2012), and donors are more willing to consider new charities and additional giving in response to large disasters (Network for Good, 2010). Donor loyalty and retention tend to be influenced by intimate and emotional relationships people have with the charities they give to (Network for Good, 2010). The time of year and even day of month also influence giving trends; in the online sphere, for example, individual giving spikes during December (accounting for approximately a third of annual giving), and within December, giving peaks in the last few days of the month. Public recognition tends to increase individual giving (Karlan & McConnell, 2012), and social pressure can be an important factor in certain types of individual giving practices (DellaVigna *et al.*, 2012).

While the charity rankers influence a large amount of private donor giving, most individuals make their giving decisions based on the aforementioned factors rather than following the recommendations of charity rankers. This is starting to change; charity ranking has rapidly grown in its influence over individual giving. GiveWell, for example, influenced US\$3million between 2007 and 2010, and in 2013 alone, its influence in giving has risen to more than US\$17million (GiveWell, 2014). The charity rankers seek to alter donor decision-making and their giving practices to align with their definition of effectiveness, and this is increasingly occurring. One survey found that 92% of people who visited Charity Navigator (the largest charity ranker) changed their giving decisions based on the information provided by it (Berger, 2010). In the case of smaller, lesser-known organisations, a high rank may result in a level of attention that would not have otherwise been achieved (Yoruk, 2013).

In order to change donor practices, the charity rankers have respectively developed a series of simple and easy to use ranking systems, such as ‘five-star ranking’ and a ‘top three list’. These rankings enhance the ‘feel good’ factor in private giving as donors are told that their donation is being given to the most effective project or organisation available to them and will have the greatest impact as a result. It is questionable, however, if the ranking systems appropriately assess effectiveness. Charities can simply adjust their reporting, not practices, to improve their rank, such as adding an informational message to a fundraising promotion and reclassify the cost from fundraising to project costs (Capin & Tanebaum, 1998). Adjusting the value of gifts-in-kind can help both donors and charities (AIP (American Institute of Philanthropy), 2009; Kristof, 2009). Of greater concern is that organisations may shift their practice entirely, to adapt to a changing donor environment that favours low-cost technocratic interventions, rather than be based upon need and local prioritisation. Donor-driven shifts in the sector occur as funding is often constrained to specific types of activities, forcing organisations to adapt to the requirements and conditionalities of donor demands.

A CASE STUDY OF MEASURING AID AND PROJECT EFFECTIVENESS

GiveWell has developed a unique ranking system, aiming to include impact when determining which charities are to be recommended and listed as top ranked. In doing so, it evaluates project types to understand which interventions have the greatest impact per dollar

spent. As such, GiveWell produces a dollar amount per life saved, and uses that amongst other metrics, including track record, room for more funding, transparency and accountability.

GiveWell focuses upon single interventions and, to an extent, organisations with a narrow programming focus. This is largely a product of the metrics and approach, as organisations implementing a cluster of interventions and/or integrated programming are more difficult to evaluate. Projects that involve multiple interventions are difficult to measure, not only due to the diversity of activities but also the range of impacts. Similarly, projects that focus upon one-time action, such as distributions of goods or health interventions, are more common as they are easier to gauge than, for example, changes in stigma or women's empowerment. In addition, public health interventions are easier to assess in its cost-per-life-saved model. Public health interventions also have easily reportable short-term impacts, which can simplify the measurement process, in contrast to policy change advocacy or good governance and capacity building projects that are complex, long term and may have diverse impacts. As a result of the available data and ease of evaluation, distribution-style projects, and primarily those within the public health sector, are ranked highly by GiveWell.

Givewell Project Evaluations: Insecticide Treated Nets

From 2009 until November 2013, GiveWell ranked the UK-based Against Malaria Foundation (AMF) as its top-rated charity, which received over US\$10million as a result. This ranking was determined after analysing not only the organisation but also the intervention type implemented by the organisation, in this case, the distribution of long-lasting insecticide-treated nets (LLINs). Although GiveWell states that it focuses on how well, in terms of implementation, programmes actually work, their analysis relies heavily on related research and existing impact evaluations, rather than actual impact assessments or project evaluations.³

In response to the question: 'Do people use the free bed nets they get?' GiveWell states this is one of their concerns. Based on studies analysed by GiveWell, uptake might be in the range of 60–80%, but no reliable or statistically significant monitoring and evaluation data are available for AMF's work. In addition, the cited studies were small-scale distributions, and AMF switched to large-scale distributions as of 2011. Furthermore, some of the cited studies included public health education and marketing, but AMF does neither. Of note is that some of the studies found that reported use differed from actual use by 12–15% (Gamble *et al.*, 2006; Lengeler, 2004). There is no reliable or statistically significant information given about the correct use of bed nets, consistent use and use over the long term for AMF's work. These distribution practices also do not take into account power dynamics or gender disparity, despite the fact that the number of bed nets distributed is not directly related to the number of women or children sleeping under these nets (Glassman & Sakuma, 2013). The cost-effectiveness agenda and charity ranker metrics often neglect the role of politics, power and gender in promoting a technocratic methodology that is based upon potential impact rather than actualized impact. Neglecting

³The information about GiveWell and AMF comes from GiveWell's website (www.givewell.org), the AMF review page (<http://www.givewell.org/international/top-charities/AMF>) and GiveWell's analysis of LLIN distributions (<http://www.givewell.org/international/technical/programs/insecticide-treated-nets>).

these important variables results in the overlooking of contextualised and tailored efforts that are better suited to the environment within which charitable activity occurs.

In Malawi, where AMF operates, the Global Fund has distributed enough nets so that each household ought to have, on average, three nets—but only 55% of households do, only 56% of children sleep under one and those who do own nets do not use them consistently (Glassman & Sakuma, 2013). Although AMF is now starting to request follow-up surveys (1–2% of households for large-scale distributions), GiveWell has found that AMF is unable to do the assessments on the required scale so the actual uptake level remains unknown. AMF also relies on third-party partner organisations to do the information gathering, adding layers of uncertainty. GiveWell itself states that the methodology, auditing and accuracy of existing usage reports are questionable.

The distribution of LLINs by the recommended charity was done using a system whereby local NGOs make proposals for nets; however, such requests were not based on needs assessments. AMF is now starting to request pre-distribution surveys to address this. As of 2011, AMF shifted to larger distributions; these too were not based on needs assessments but information on probable coverage gaps in the national programme. The current method of monitoring includes taking photos and video of distributions, so with confidence, GiveWell can say that LLINs are being given out. This raises the following questions: (i) To whom are these nets given, as the distributions are not based on needs assessments? (ii) Are they being used consistently and correctly? (iii) What is the likelihood of continued use over the long term? (iv) Because of the limited shelf life of LLINs, and a lack of re-distribution plans, how is AMF ensuring the impact is sustained? The answers to these questions lie beyond the realm of GiveWell's metrics, and therefore, the answers are unknown.

A team responsible for calculating the cost-effectiveness of a programme, such as LLIN bed net distribution, may not pose the previous questions without having experience in the practice of international development. This is one of the key challenges facing GiveWell, as none of the full-time staff have a background in international development or humanitarian work. Over 60 years of development aid have demonstrated that complex challenges in developing nations, such as extreme poverty, are not easy to resolve or evaluate. In fact, they are often embedded in complex and poorly understood environments. The notion that a technocratic solution can be handed down simply and cheaply is, perhaps, reflective of how a ranking service understands the context within which their recommended organisations operate.

GiveWell knows that LLIN distribution can be effective, based on the theory and evaluations conducted of other interventions, but in practice, little is known about the top-ranked charity's actual impact. GiveWell suggests that applying the results of studies from one area to another is sufficient to support their recommendations of the best giving opportunities, along with their other criteria. However, the data from other countries do not speak to the context of Malawi, or the specific context wherein distributions are taking place, particularly, as mentioned before, where large-scale distributions in Malawi have not been translated into high uptake and usage, or equitable coverage (Glassman & Sakuma, 2013). Therefore, it is not clear why GiveWell believes that more large-scale distributions will address these particular and more complex challenges. Perhaps improving programme integration, for example, including public health education, environmental interventions to reduce mosquito breeding grounds, and diagnosis and treatment, is required. This, however, would increase the per person cost effectiveness of the interventions. The Global Fund, which also works in Malawi, funds programming in

this comprehensive fashion, and is working towards elimination of malaria in 32 countries and engaging in malaria control in another 67 countries (Global Fund, 2013). The elimination of malaria is a permanent interruption of mosquito-borne malaria transmission, an achievement that LLIN distribution cannot achieve as a stand-alone intervention. This also calls into question the duration of ‘effectiveness’, as considered by the charity rankers.

Other considerations of this approach also require greater reflection. We might consider asking: ‘To what extent ought an intervention be considered in isolation of the rest of society?’ The model implemented by GiveWell’s top-ranked charity, sourcing LLINs externally at a lower cost, as opposed to purchasing them from local or regional manufacturers may affect the long-term ability of nations to have the capacity to respond to malaria domestically. Instead, the approach emphasises lowest cost, and therefore imports the cheapest goods; this may facilitate a situation where, in the short term, donors purchase low-cost LLINs while the capacity of the nation to ensure domestic supply of LLINs is reduced. Of note, the top-ranked charity has no plan to redistribute LLINs, which is important because LLINs have a life span of 2–5 years. As it provides no public health education, this may be unknown to recipients of the LLINs. Being an area where malaria is endemic, short-term hand-outs will not build or support the local or national capacity to have LLIN coverage or reduce malaria in the long term. Locally and nationally driven societal change, including those driven by external funding, can have multiple impacts beyond the direct ones. For example, sourcing LLINs within Malawi could have been economically beneficial for the country by developing the manufacturing capacity for the long term, enhancing practical skills, supporting local employment and improving livelihoods in the process. Shifts of this nature have occurred for food aid, which also improved timeliness and cost-effectiveness (Tschirley & del Castillo, 2006; Lentz *et al.*, 2012). Much more reflection is required regarding how charitable systems can make systematic shifts to enhance impact in multiple spheres while maintaining effectiveness. Far too little reflection of this nature occurs within the charity-ranking sphere.

GiveWell does not discuss the potential effectiveness of bundling interventions nor does it evaluate other organisations operating in the same region that do bundle interventions. Bundled interventions are more challenging to measure than singular interventions. However, other rankers that have evaluated bundled interventions, such as the Copenhagen Consensus, have found bundled interventions to be more effective than singular ones. An outcome of this type of ranking is that effectiveness fits within a certain set of parameters.

While this analysis has considered the context of one organisation, it is the way in which GiveWell evaluates this organisation that we wished to highlight. In order to sufficiently assess the metrics and evaluation processes of GiveWell, a sufficient amount of detail was required regarding its top-ranked charity. The theoretical lessons learned from the GiveWell study raises questions regarding how charity rankers conceptualise charitable activity, apply metrics and, ultimately, the ways in which it ranks a charity.

Missing Metrics

Metrics are the tools used in evaluation. Although rarely the subject of much attention, the choice of metrics greatly affects the way in which evaluation takes place, what the evaluation can and cannot tell as well as places an emphasis upon certain components over others. The charity rankers all use metrics in their evaluations, and these processes

influence the results. This section outlines some of the missing metrics and resulting impact thereof.

The trend in ranking charities is that organisations and interventions are understood in the context of econometrics and cost-effectiveness. This approach has benefits and drawbacks. As the aid sector is a complex environment for donors to navigate, a financial evaluation makes the ranking process possible, and it establishes metrics wherein the complexity of aid can be understood, measured and ranked. It also allows for an investigation into what the costs of respective interventions are and to compare those costs with other interventions.

Whereas costing mechanisms are often drawn from similar evaluations, these tend to be specific to the unique context within which they took place. An RCT that took place in Ecuador may be applied to the costing calculations for programming in Malawi, for example. However, the sociocultural, political, environmental and disease pattern dynamics differ significantly. In addition, the partners with whom organisations operate in order to conduct experiments greatly influences the outcomes, and make replication and scale-up far more questionable and challenging than assumed or implied (Allcott & Mullainathan, 2012).

Econometric-based evaluations do not take into account human rights, and human rights do not appear as metrics in determining what is, and is not, effective by charity rankers. Some research (Kabeer, 2010) has shown that target-based objectives may further entrench the disempowerment and discrimination of certain segments of society. A process of decision-making based upon cost effectiveness is one that is not primarily focused on human rights, which means that, rather than those who are in greatest need, those who are the most cost-efficient to reach may be targeted. In such a case, the easiest to reach and/or largest potential intervention areas are targeted to reduce costs, while minority and disadvantaged individuals and groups may be further marginalised. An example of this is an intervention reaching people living near a paved road, as this reduces projects costs, while those living further from transportation infrastructure are not, as this would increase the per-person cost. Chambers (2006: 17) has called this the 'tarmac and roadside biases' and it is often the case that those in the rural and remote areas are in greater need of this intervention and have more limited access to the provided goods, such as LLINs.

In addition to affecting the selection of beneficiaries, the intervention types selected by charity rankers are those focused on increasing cost effectiveness. The case of mosquito net distribution, highlighted above, exemplifies how a cost-effective method distributes LLINs in bulk and video records distribution events as a monitoring mechanism. A programme focused on achieving the stated objective might include a needs assessment, targeted distribution and information for beneficiaries on how and when to best use nets, provision of public health education, follow-up and plans for regular redistribution upon estimated expiry. As this comes with significant additional costs, the dollar-per-life-saved goes up, negatively affecting the ranking process and therefore discourages this intervention approach from taking a long-term and system-wide perspective.

The charity rankers we evaluated, most of which were not explored in this work, do not take into account the context where projects operate, and tend to promote a top-down and technocratic model of development. Although aspects of community-driven development, such as empowerment and local agency, have been challenged by some (Duflo, 2012), not all projects fall under one category. Rules and regulations regarding health and safety, for example, wearing a seatbelt or determining dangerous levels of contaminant content in foods, ought to be administered at the national and international level. As such, it is not

argued that all action ought to be bottom-up and participatory. However, certain forms of charitable activity should as it lends to greater effectiveness and sustained change (Lawry, 2008) and can address important issues of equity (Harper *et al.*, 2014). Development practitioner and public policy professor, Krishna, writes:

Knowing contextually-relevant reasons for escape and descent [into poverty] is a pre-condition for effective program design, but rarely have external assistance programs identified and targeted context-specific reasons. In the most part, programs of assistance have followed a pre-determined and broad-brush logic, one that is dictated by people far removed from (and often ignorant about) diverse and changeable local conditions (Krishna, 2010: 108).

External players can be instrumental in setting up systems, and Duflo (2012) provides examples of clean water systems and retirement savings plans. Fasil Kiros, an Ethiopian economist, states that if one important lesson is to be learned from aid relationships and donor-driven policy, 'it is the fact that none of the international donors who usually come up with new development paradigms have a better knowledge of the needs and development potentials' than nations themselves (Kiros, 2006, 115). The basis of his claim is that donors rapidly abandon advocated policies for new ones before the effects of the former have time to develop and be evaluated. Policy-driven decision-making may be suitable for certain circumstances; however, many interventions evaluated and recommended by the charity rankers do not fall in this category. As such, contextualised programming ought to be key to the determination of effectiveness.

In addition, advocating particular low-cost high-impact interventions does not mean they will be welcomed. Carothers and de Gramont (2013: 249) find that, at the national level, 'even if such studies [RCTs] are able to come up with some generalizable truths, their focus on evidence and on identifying effective policies leaves out the crucial question of whether host governments will be willing and able to implement these recommendations'; the same applies at the local level. An example of this is the work of international advocacy groups seeking to stop female genital mutilation/circumcision in developing countries, which has resulted in backlash of protests rejecting and opposing these 'development' projects (Ruderman, 2013).

A few examples of how local priorities and appropriateness matter are in order. An area with endemic malaria needs LLINs, but amidst chronic food insecurity, its distribution may not be met with enthusiasm and uptake may remain low as other needs are prioritised. While the theoretical data is correct, it does not match the context. The removal of landmines to enable farming, the creation of a centralised food storage system to ensure supplies last through the dry season and the drilling of water bores when the river is a source of cholera may all be prioritised higher, in their respective contexts, than gaining access to bed nets. Contextualised and appropriate interventions that are based upon needs assessments affect impact in practice and must be incorporated into evaluations for effectiveness to be more appropriately understood.

UNINTENDED OUTCOMES

Interventions deemed to be most cost effective do not commonly build the capacity of permanent stakeholders, such as the institutions and personnel in the public healthcare system, to improve the sustainability of the intervention. As Carothers and de Gramont (2013: 6) argue 'without minimally effective government institutions to support

sustainable development, progress is likely to remain out of reach for many countries'. Beyond those basic functions, increased state capacity has the potential to improve provision of goods, as well as prosperity over the long term (Acemoglu *et al.*, 2013). The impact of such interventions seems not to be weighted in costing calculations, which are based on a short-term timeframe.

As the cost-effectiveness approach gains momentum, it will have a greater impact upon organisations. Rather than continue work within an area of expertise in response to local needs, organisations may shift into low-cost interventions to secure funding—from advocacy of gender equality to micronutrient supplement distributions, for example. The power inequality between donor and recipient organisation cannot be underestimated and organisations will undoubtedly adapt to changes in donor giving. This lessens the ability for beneficiaries to participate in programming to help shape interventions that meet the pressing needs of their respective contexts. Criteria of organisational accountability and evidence-based uses of donor funding are not harmful in and of themselves; however, when metrics are used in isolation or used in ways that exclude other important methods of assessment, unintended consequences can result.

Robert Chambers (2012) argues that development aid gets stuck in a 'lock-in', which he calls a paradigmatic syndrome, wherein there is strong mutually-supporting inflexibility such that alternatives beyond the current paradigm are not considered. In other words, practitioners and donors cannot 'think outside the box' or see any other possible solutions. He gives examples of this, such as development interventions seeking to improve hygiene and sanitation. Once someone breaks out of that 'lock-in', it can have major impacts; Community-Led Total Sanitation, an innovative community-driven methodology for mobilising communities to eliminate open defecation, is an example of that and has now affected more than 10 million people. Prior to that innovation, the aid community was stuck in a programming mode that advocated a similar set of interventions that were not really working (or marginally so). The break of a 'lock-in' might be counterintuitive, as the sanitation example turns 'best practice' upside-down with relatively simple, yet unasked, questions. In our assessment, most charity rankers are locked-in to a specific model of ranking, focusing upon specific metrics, its definition of cost-effectiveness and assumptions about what impacts should occur and how.

An Alternative?

Most charity rankers have not broken out of the 'lock-in' of assessments and do not provide alternatives to the technocratic replication of cost-effective results-based projects. In seeking to encourage the improvement of giving practices and recognising the potential of simple and easy to use models to support that, we do not suggest the entire enterprise of charity ranking be abandoned. However, as with Krishna, we believe a radical shift is required:

Identifying, and then targeting, context-specific negative and positive reasons [of entering and exiting poverty] needs to become the guiding principle for assistance programs of the future. There is no magic remedy that will be equally efficacious in all parts of the developing world. Standardized programs extended across large numbers of regions and countries will very likely continue to have relatively little positive effect. Programs that are designed after gaining knowledge of locally

operating threats and opportunities will more effectively tilt the balance of events in the right direction (Krishna, 2010: 108).

As mentioned at the outset, one of the objectives of this paper was to contribute to a discussion and provoke critical reflection about current trends, rather than propose a new model. Charity rankers are limited by their access to information, constrained by human and resource capacity and cannot practically assess each and every local priority around the world. The answer may not lie in changing metrics while using the same method—it may require a different approach altogether. GlobalGiving is an example of an existing model seeking to influence individual giving practices that better reflects Krishna's direction for a more appropriate, contextualised and locally prioritised funding model. In this case, grassroots projects are proposed, screened and implemented by local organisations. Anyone can nominate a project, followed by an application, screening and fundraising trial. Local community groups can propose small to mid-sized projects that address their needs and priorities, rather than being recipients of pre-determined programming and costing calculations made on other continents. Although it does not address all the concerns raised in this paper, and indeed raises concerns of its own, it provides an example of an innovative model for the charity-ranking sector wherein community members are the decision makers and implementers and has the flexibility to support a wide array of potential new, and innovative, ideas. We hope that critical discourses such as this one will facilitate the emergence of new ideas, models and approaches that take into account the concerns that have been raised, and others that have yet to be identified, so as to move towards a giving environment that is increasingly reflexive of its intended and unintended impacts.

CONCLUSIONS AND RECOMMENDATIONS

While the approach taken by charity-ranking bodies is not uniform, the focus upon costing and the neglect of contextualization in an effort to create simple and generalizable assessments can negatively impact the objective of charity ranking and potentially counter the promotion of human rights-based efforts. The metrics used in identifying 'cost-effectiveness' may be missing important components and, as a result, create a set of unintended outcomes that negatively affect beneficiaries, the humanitarian and development sector and, importantly for this discussion, the effectiveness of charitable activities.

The methodologies developed and advocated by the charity rankers are a reflection of broader trends occurring within development. The neoliberalisation and professionalisation of the sector, while unique in their respective manifestations, share an ideology that promotes technocratic and calculating approaches. Equity and human rights are challenging to measure; however, a significant amount of data is available, such as in the annual Human Development Reports produced by the United Nations Development Programme (2015). A lack of currently available data ought not to prevent thought on how new and better data may be obtained. It is noteworthy that despite the existence of some data, and the availability of potential indicators, the charity rankers do not incorporate human rights into their models. We, therefore, suggest that the focus on cost-effectiveness must also incorporate broader analyses that include greater contextualization and an inclusion of human rights to ensure that funding does not foster greater inequality and entrench marginalisation but acts as an effective means to improve human rights and equity.

REFERENCES

- Acemoglu D, Garcia-Jimeno C, Robinson J. 2013. State capacity and economic development: a network approach. Draft 16 September 2013.
- AIP (American Institute of Philanthropy). 2009. Charity questions the value of donated goods [online]. Available at: <<http://www.charitywatch.org/articles/feedchildren.html#value>> [Accessed 3 April 2014].
- Allcott H, Mullainathan S. 2012. External validity and partner selection bias [online]. Available at: <<http://www.nber.org/papers/w18373>> [Accessed 11 October 2013].
- Andrews M. 2013. *The Limits of Institutional Reform in Development: Changing Rules for Realistic Solutions*. Cambridge University Press: Cambridge.
- Barnes T. 2001. Rethorizing economic geography: from the quantitative revolution to the “cultural turn”. *Annals of the Association of American Geographers* **91**(3): 546–565.
- Berger K. 2010. Charity navigator: financial measures and beyond. Presentation at InsideNGO: Finance Grants and Contracts Annual Meeting, 13 July 2010.
- Breeze B. 2010. How Donors Choose Charities: Findings of a Study of Donor Perceptions of the Nature and Distribution of Charitable Benefit. Centre for Charitable Giving and Philanthropy: London.
- Buffardi A. 2011. Incorporating donor dimensions into aid implementation and effectiveness research. *Evans School Review* **1**: 75–91.
- Capin G, Tanebaum J. 1998. How to Report a Joint Activity: With new accounting rules in place, NPOs may need to change how they report the results of fundraising. *Journal of Accountancy*, July 31. Available at: <<http://www.journalofaccountancy.com/issues/1998/aug/capin.html>> [Accessed 7 November 2015]
- Carothers T, de Gramont D. 2013. *The Almost Revolution: Development Aid Confronts Politics*. Washington: Carnegie Endowment.
- Chambers R. 2012. *Provocations for Development*. Practical Action Publishing: Warwickshire.
- Chambers R. 2006. *Poverty Unperceived: Traps, Biases and Agenda*. Institute of Development Studies: Brighton.
- Collier P, Dollar D. 2001. Can the world cut poverty in half? How policy reform and effective aid can meet international development goals. *World Development* **29**(11): 1787–1802.
- Deaton A. 2010. Instruments, randomization, and learning about development. *Journal of Economic Literature* **48**(2): 424–455.
- DellaVigna S, List J, Malmendier U. 2012. Testing for altruism and social pressure in charitable giving. *The Quarterly Journal of Economics* **127**(1): 1–56.
- Dollar D, Pritchett L. 1998. Assessing aid – what works, what doesn’t, and why. World Bank policy research report. Washington, D.C.: The World Bank. <http://documents.worldbank.org/curated/en/1998/11/438890/assessing-aid-works-doesnt> (Accessed 1 April 2014).
- Duflo E. 2012. Human values and the design of the fight against poverty [online]. Available at: <economics.mit.edu/files/7904> [Accessed 3 December 2013].
- Easterly W, Williamson C. 2011. Rhetoric versus reality: the best and worst of aid agency practices. *World Development* **39**(11): 1930–1949.
- Ebrahim A. 2010. The many faces of nonprofit accountability. Harvard Business School, Working Paper 10–069.
- Eyben R. 2000. Development and anthropology: a view from inside the agency. *Critique of Anthropology* **20**: 7–14.
- Eyben R. 2011. In *The Sociality of International Aid and Policy Convergence*, in *Adventures in Aidland: The Anthropology of Professionals in International Development*, Mosse D (ed). Berghahn Books: New York.

- Fuchs A, Dreher A, Nunnenkamp P. 2014. Determinants of donor generosity: a survey of the aid budget literature. *World Development* **56**: 172–199.
- Gamble C, Ekwaru J. P., ter Kuile F. O. 2006. Insecticide-treated nets for preventing malaria in pregnancy. *Cochrane Database Syst Rev*, **2**(2).
- GHA. 2012. Global Humanitarian Assistance Report 2012. Global Humanitarian Assistance: Somerset.
- GiveWell. 2014. Evaluating GiveWell's impact [online]. Available at: <<http://www.givewell.org/about/impact>> [Accessed 23 November 2014].
- Glassman A, Sakuma Y. 2013. How many bed nets equal a life saved? – Why results matter for value for money [online]. Available at: <<http://www.cgdev.org/blog/how-many-bed-nets-equal-life-saved-%E2%80%93-why-results-matter-value-money>> [Accessed 11 October 2013].
- G Fund. 2013. Malaria [online]. Available at <<http://www.theglobalfund.org/en/about/diseases/malaria/>> [Accessed 12 December 2013].
- Gould J. 2005. *The new Conditionality: The Politics of Poverty Reduction Strategies*. Zed Books: London.
- Gulrajani N. 2011. Transcending the great foreign aid debate: managerialism, radicalism and the search for aid effectiveness. *Third World Quarterly* **32**(2): 199–216.
- Harper C, Nowacka K, Alder H, Ferrant G. 2014. *Measuring Women's Empowerment and Social Transformation in the Post-2015 Agenda*. Overseas Development Institute: London.
- Harvey D. 2003. *The New Imperialism*. Oxford University Press: Oxford.
- Harvey D. 1982. *The Limits to Capital*. Blackwell: Oxford.
- Heiberg D. 2009. Which measurement matters? [online]. Available at: <https://www.maxwell.syr.edu/uploadedFiles/moynihan/tngo/MD_Measurement_Sept09.pdf> [Accessed 13 November 2013].
- Hermes N, Lensink R. 2001. Changing the conditions for development aid: a new paradigm? *Journal of Development Studies* **37**(6): 1–16.
- Jones D. 2011. Information Revelation in a Public Goods Game: An Experimental Study of the Role of Charity Ratings Organizations. AccessEcon Submission No. PET11-11-00299. Available at: <<http://www.accessecon.com/pubs/PET11/GeneralPDFPET11/PET11-11-00299S.pdf>> [Accessed 19 November 2015].
- Kabeer N. 2010. *Can the MDGs Provide a Pathway to Social Justice? The Challenge of Intersecting Inequalities*. New York: United Nations Development Programme.
- Karlan D, McConnell M. 2012. Hey look at me: the effect of giving circles on giving.
- Kiros F. 2006. *Enough with Famines in Ethiopia: A Clarion Call*. Tsehai Publishers: Hollywood.
- Knack S, Rogers F, Eubank N. 2011. Aid quality and donor rankings. *World Development* **39**(11): 1907–1917.
- Kothari U. 2005. Authority and expertise: the professionalisation of international development and the ordering of dissent. *Antipode* **37**(3): 425–446.
- Krishna A. 2010. *One Illness Away*. Oxford University Press: New York.
- Kristof K. 2009. Charities fake their numbers to look good [online]. Available at <<http://www.cbsnews.com/news/charities-fake-their-numbers-to-look-good/>> [Accessed 3 April 2014].
- Lawry S. 2008. Early Ford foundation support for the Grameen Bank: Lessons in philanthropic accountability, risk, and impact. Harvard University, Working Paper No. 44.
- Lengeler C. 2004. Insecticide-treated bed nets and curtains for preventing malaria. *Cochrane Database of Systematic Reviews*, Issue 2. Art. No.: CD000363. DOI:10.1002/14651858.CD000363.pub2.
- Lentz E, Passarelli S, Barrett C. 2012. The timeliness and cost-effectiveness of the local and regional procurement of food aid [online]. Available at: <[http://dyson.cornell.edu/faculty_sites/cbb2/Papers/LRP%20Ch%20%20Lentz%20et%20al%20Jan%2011%202012%20\(1\).pdf](http://dyson.cornell.edu/faculty_sites/cbb2/Papers/LRP%20Ch%20%20Lentz%20et%20al%20Jan%2011%202012%20(1).pdf)> [Accessed 03 December 2013].
- Lewis A. 1955. *The Theory of Economic Growth*. Unwin Hyman: London.

- Li TM. 2007. *The Will to Improve: Governmentality, Development, and the Practice of Politics*. Duke University Press: Durham.
- Lowell S, Trelstad B, Meehan B. 2005. *The Ratings Game*. Stanford Social Innovation Review: Stanford.
- Mathenge P. 2013. Challenges facing project implementation in selected public sector organizations in Kenya. School of business, Kenyatta University, D53/CTY/PT/23270/2011.
- McGillivray M. 2003. Aid effectiveness and selectivity: integrating multiple objectives into aid allocations, WIDER Discussion Papers, World Institute for Development Economics (UNU-WIDER), No. 2003/71.
- McGillivray M, Feeny S, Hermes N, Lensink R. 2006. Controversies over the impact of development aid: it works; it doesn't; it can, but that depends. *Journal of International Development* **18**(7): 1031–1050.
- Mekasha T, Tarp F. 2013. Aid and growth: what meta-analysis reveals. *Journal of Development Studies* **49**(4): 564–583.
- Millikan MF, Rostow WW. 1957. *A Proposal: Key to an Effective Foreign Policy*. Harper: New York.
- Mitchell GE. 2014. Creating a philanthropic marketplace through accounting, disclosure, and intermediation. *Public Performance & Management Review* **38**: 23–47.
- Mosse D. 2004. Is good policy unimplementable? Reflections on the ethnography of aid policy and practice. *Development and Change* **35**(4): 639–671.
- Mosse D. 2010. A relational approach to durable poverty, inequality and power. *Journal of Development Studies* **46**(7): 1156–1178.
- Murray W, Overton J. 2011. Neoliberalism is dead, long live neoliberalism? Neostructuralism and the international aid regime of the 2000s. *Progress in Development Studies* **11**(4): 307–319.
- Network for Good. 2010. The online giving study [online]. Available at: <http://www.fundraising123.org/files/Community/Online_Giving_Study_2010R.pdf> [Accessed 2 December 2013].
- Overton J, Murray W, McGregor A. 2013. Geographies of aid: a critical research agenda. *Geography Compass* **7**/2(2013): 116–127.
- Pearson L. 1969. *Partners in Development*. Praeger: New York.
- Rosenstein-Roden PN. 1943. Problems of industrialisation of Europe and south-eastern Europe. *The Economic Journal* **53**(21): 202–211.
- Rose N. 1999. *Powers of Freedom: Reframing Political Thought*. Cambridge University Press: Cambridge.
- Ruderman R. 2013. Female circumcision: the ethics of harm reduction policies. *The Michigan Journal of Public Affairs* **10**: 95–107.
- Singer P. 2009. *The Life You Can Save: Acting Now to End World Poverty*. Random House: New York.
- Tschirley D, del Castillo AM. 2006. Local and regional food aid procurement: an assessment of experience in Africa and elements of good donor practice. *Policy Synthesis* **79**: 1–4.
- United Nations Development Programme. 2015. About human development. Available at: <http://hdr.undp.org/en/humandev> [Accessed 4 April 2015].
- Yoruk B. 2013. Charity ratings, working paper [online]. Available at: <<http://www.albany.edu/economics/research/workingp/2013/yoruk3.pdf>> [Accessed 13 November 2013].