

Program Evaluation and Impact Assessment in International Non-Governmental Organizations (INGOs): Exploring Roles, Benefits, and Challenges

Stan Yu

University of Saskatchewan

Darrell Arthur McLaughlin

St. Thomas More College, University of Saskatchewan

ABSTRACT

In the twenty-first century, the call for International Non-governmental Organizations (INGOs) to demonstrate their effectiveness has become popularized. This has given rise to scholarly attention examining the roles of program evaluation and impact assessment in assisting INGOs in demonstrating their effectiveness. While previous studies suggest that INGOs actively conduct program evaluation and impact assessment, this article explores the perspectives of two Canadian INGOs on how they understand, use, and experience evaluation and assessment as it relates to their work. Our study uncovers three continuing challenges: evaluation and assessment are largely descriptive and lack more sophisticated analyses; efforts to conduct evaluation and assessment are consolidated within organizations' head offices, while staff members and volunteers are largely excluded; and evaluation and assessment remain rooted in the paradigm of quantifiable results, which do not truly reflect the nature of work being conducted on the ground.

RÉSUMÉ

Au vingt-et-unième siècle, on veut de plus en plus que les organisations non gouvernementales internationales (ONGI) démontrent leur efficacité. Ce désir a motivé les chercheurs à se pencher sur les évaluations de programme et les études d'impact pour voir dans quelle mesure celles-ci peuvent aider les ONGI à montrer qu'elles sont efficaces. Des études antérieures suggèrent que les ONGI mènent de manière concertée des évaluations de programme et des études d'impact. Cet article explore comment aujourd'hui deux ONGI canadiens comprennent, utilisent et vivent l'évaluation et la mesure de leur travail. Notre étude relève trois défis actuels : l'évaluation et la mesure tendent à être descriptives sans offrir d'analyses plus poussées; ce sont les sièges sociaux des organismes qui gèrent l'évaluation et la mesure en excluant ainsi bon nombre de fonctionnaires et volontaires; l'évaluation et la mesure se limitent au paradigme des résultats mesurables et par conséquent elles ne reflètent pas nécessairement le véritable travail mené sur le terrain.

Keywords / Mots clés : Non-governmental Organizations; Program evaluation; Impact assessment; Effectiveness; Qualitative research; International development / Organisation non gouvernementale internationale; Évaluation de programme; Étude d'impact; Efficacité; Recherche empirique; Développement international

INTRODUCTION

Since the late 20th Century, the call for International Non-governmental Organizations (INGOs) to demonstrate their effectiveness has become increasingly popularized (Abdel-Kader & Wadongo, 2011; Ebrahim & Rangan, 2010; Lecy, Schmitz, & Swedlund, 2011; Morley, Vinson, & Hatry, 2001; Moxham, 2009; Spar & Dail, 2002). This position is epitomized by Fisher (1997) who explains that in the 21st Century, the untainted image of INGOs as “doing good” to provide the world with “the service of a social need neglected by the politics of the State and the greed of the market,” was suddenly met with the sobering reality that after thirty years of increased numbers, budgets, and responsibilities, INGOs had yet to show the world any substantial change (Atack, 1999). Edwards and Hulme (1996) further argue that the ascription of INGOs as the “magic bullet” for solving global issues often carries very little evidence to support it. For INGOs, this need to demonstrate effectiveness is intimately linked to parallel discussions of accountability and legitimacy, both of which have also been challenged by politicians, academics, the media, and the public alike (Atack, 1999; Gibelman & Gelman, 2004; Harsh, Mbatia, & Shrum, 2010; Nicolau & Simaens, 2009).

The prominence of these debates has given rise to a stream of scholarly attention examining the roles of program evaluation and impact assessment within INGO work (Alaimo, 2008; Bouchard, 2009; Campbell, 2002; Ebrahim & Rangan, 2010; Moxham, 2009). For these scholars, the focus placed on the role of program evaluation and impact assessment comes with both promise and concern. On the one hand, the use of program evaluation and/or impact assessment offers a critical tool for INGOs to respond to the challenges of effectiveness by providing empirical evidence of their impact with some degree of academic standard, rigour, and objectivity, while also cultivating organizational learning and best practices within INGOs themselves (Alaimo, 2008; Campbell, 2002; Edwards, 1997; Travers, 2011). On the other hand, some suggest that program evaluation and impact assessment contribute to furthering the neoliberal “new public management” (NPM) model, or the increasing adoption of business management principles, such as efficiency, competition, entrepreneurship, consumer-driven, and a focus on measuring performance and outcomes in the nonprofit sector (Kilby, 2004; Sarker, 2005).

Despite the substantial scholarly attention that has been placed on the roles of program evaluation and impact assessment in INGOs, many areas of discussion remain in need of further examination, including how to define, operationalize, and measure indicators that accurately capture INGO performance and impact; identifying suitable measurement systems for the work of INGOs; standardizing the social indicators of effectiveness for INGOs across diverse fields of nonprofit work and resolving whether that is even a worthy goal to strive for; and determining to what extent program evaluation and/or impact assessment are being conducted within INGOs (Abdel-Kader & Wadongo, 2011; Bouchard, 2009; Lecy, Schmitz, & Swedlund, 2011). We argue that the final topic of discussion is of fundamental importance as it establishes a foundation through which the other discussions can occur. Specifically, we suggest that questions of capacity and capability for conducting program evaluation and impact assessment, as well as what role(s) findings and results play after conducting program evaluation and impact assessment within INGOs, need to be more closely examined. While the literature has described the ideal role that program evaluation and impact assessment should play within INGO work, our interest lies in how that ideal compares to how INGOs actually understand and use program evaluation and impact assessment on the ground.

This article contributes to addressing this question. Specifically, this article explores the perspectives of two Canadian INGOs working in international development, and how their staff members and volunteers understand and experience program evaluation and impact assessment as it relates to their work. More specifically, we ask questions such as: Are program evaluation and/or impact assessment being conducted within the two organizations? Who is primarily responsible for conducting these evaluations and/or assessments? What role does program evaluation and/or impact assessment play within the organization? How are findings and results reported and disseminated to internal and external audiences? How are findings used within existing programs and interventions within the organization?

In turn, this article is made up of three main sections. First, a review of the current literature on program evaluation and impact assessment in INGO work is presented, underscoring the benefits and challenges of conducting evaluation and assessment in INGO work. As well, we report on findings from previous studies that have investigated the extent to which INGOs are actively conducting evaluation and assessment. Second, the methodological procedures of this study are described and a profile of the two participating organizations is provided. Finally, the findings of the study are reported, their implications for the literature are discussed, and recommendations for future studies in this area are detailed.

PROGRAM EVALUATION AND IMPACT ASSESSMENT IN INGO WORK

While there exists an intimate linkage and a significant degree of overlap between the concepts of program evaluation and impact assessment, there are also fundamental differences between the two terms, warranting each to be a unique and exclusive term within the wider INGO scholarship. First, program evaluation can be defined as a systematic process of collecting information and applying approaches, techniques, and knowledge to analyze, research, and assess the activities, characteristics, and outcomes of a program(s) to document its accomplishments and to improve the planning, implementation, and effectiveness of these programs (Alaimo, 2008; Perret, 2009). In INGO work, program evaluation can be conducted before, during, or after the implementation of a program or intervention (Bouchard, 2009; Campos, Andion, Serva, Rossetto, & Assumpção, 2010).

Meanwhile, impact assessment can be defined as the systematic process of analyzing significant changes, whether positive or negative, intended or unintended, as a result of a particular planned activity, program, intervention, or project, on people's lives (Dawson, 2010; Gosling & Edwards, 2003; Lockie, 2001; Moxham, 2009). Impact assessment should not be treated as mutually exclusive from program evaluation, but rather as an accession of program evaluation that focuses on the longer-term and wider-ranging changes beyond the immediate results of INGO work (Gosling & Edwards, 2003). In other words, impact assessment takes the central focus away from the program or intervention itself and examines the implications and effects of those programs on the recipients of those programs. At its most fundamental level, impact assessment is about understanding change and the key processes that led to that change (Gosling & Edwards, 2003).

The use of program evaluation and impact assessment serves to benefit INGO operations on multiple levels. First, at the programmatic level, the findings from program evaluation and impact assessment can be used by directors, managers, staff, and volunteers to assess the process, quality, efficiency, and productivity of their programs, which will inform future ground-level decisions within those programs and work toward cultivating best practices (Alaimo, 2008; Bouchard, 2009; Gosling & Edwards, 2003; Nicolau & Simaens, 2009). Second, at the organizational level, members in leadership roles at INGOs can utilize the findings to disseminate and communicate the progress of individual projects or of the organization as a whole to their

stakeholders for strategic planning and decision-making processes (Alaimo, 2008). Finally, at the societal level, INGOs can utilize the findings as evidence to contribute to the larger discussion of their effectiveness on the alleviation or minimization of the social problems they set out to address (Alaimo, 2008; Ebrahim & Rangan, 2010; Nicolau & Simaens, 2009).

While the use of program evaluation and impact assessment has been promoted within the literature as having the potential to support and further the work of INGOs, a plethora of challenges and barriers have also been documented. First, unlike for-profit organizations where success is ultimately determined by increased revenue or shareholder value, any bottom line success measures pertaining to nonprofit work (e.g., providing clean water, or improving conditions in aspects of health, economic structures, education, quality of life, etc.) pose a more difficult challenge for efforts to conduct program evaluation and / or impact assessment (Edwards & Hulme, 1996; Spar & Dail, 2002; Walsh & Lenihan, 2006).

Second, INGOs work in natural, open, unstable, and complex systems, which are subject to rapid change from the external environment (Edwards & Hulme, 1996; Fowler, 1997). This means that any observed impact is likely to be attributed to multiple factors and actors, rather than solely through the efforts of INGOs (Ebrahim & Rangan, 2010; Fowler, 1996). Third, there are institutional barriers that hamper the development of a strong focus on conducting program evaluation and/or impact assessment within INGOs, such as the internal culture of INGOs that values action more than reflection, the limitations of the instruments and expertise INGOs use to evaluate their own programs, the lack of financial flexibility to hire a specialist in the area of evaluation and assessment, and the fact that rigorous evaluation and assessment are simply not required by some funders (Barber & Bowie, 2008; Campbell, 2002; Edwards & Hulme, 1996; Fowler, 1997; Lecy, Schmitz, & Swedlund, 2011).

Lastly, some scholars question whether the emphasis on program evaluation and impact assessment within the work of INGOs is merely another symptom of donor-centric accountability within the aforementioned “new performance management” (NPM) model (Abdel-Kader & Wadongo, 2011; Agg, 2006; Brown & Moore, 2001; Serva, Andion, Campos, & Onozato, 2009). More specifically, these scholars assert that INGOs are experiencing an increasing financial dependence on funding from short-term project-based “grants” from governments and foundations, such as the Canadian International Development Agency (CIDA) and the Department for International Development (DFID) in the United Kingdom (Agg, 2006; Nicolau & Simaens, 2009; Szporluk, 2009). Thus, INGOs accepting a grant from CIDA or DFID will be confined to operate within restricted activities as specified by the donor’s grant (Agg, 2006; Parks, 2008). Catherine Agg (2006) cites the example of DFID where, in order for INGOs to receive a partnership grant, they “have to prove that they contribute to the UK government’s own ‘targets and priorities’” (p. 19). In addition, these grants exert pressure on INGOs to abandon long-term projects in international development and reframe them into more short-term and measureable projects in order to fit within the NPM framework (Brown & Moore, 2001; Nicolau & Simaens, 2009; Smith, 2008; Szporluk, 2009). The effect of this current model of funding on the quality of international development work continues to be a topic of discussion within the literature (Agg, 2006).

A final topic of interest within the literature is how common is conducting program evaluation and impact assessment within INGOs. Recent studies suggest that INGOs are actively conducting program evaluation and/or impact assessment within their work. In a nationwide survey of 162 Canadian INGOs focused on international development, Travers (2011) found that 62 of 69 (90%) responding organizations conducted program evaluation as part of their work. Morley, Vinson, and Hatry (2001) further reported that 83% of 36 nonprofit organizations regularly collected, tabulated, and examined data as part of their work. In 2010, Campos et al. similarly found that 87% (54 of 62 Brazilian NGOs) perform program evaluation.

METHODOLOGY

While previous findings have established that a large proportion of INGOs conduct some form of evaluation and assessment, it is our position that what is missing within the existing literature are the perspectives and experiences of conducting program evaluation and impact assessment from the INGOs themselves. In response, this study sets to uncover the voices of INGOs and those working within INGOs in how they understand and experience the notion of conducting program evaluation and/or impact assessment within their own work. Of interest in this study were questions that inquired as to who conducted program evaluation for their organization, what purpose(s) do the results of program evaluation serve, what are successes and/or barriers to conducting evaluation and assessment, and how are results disseminated to stakeholders, both internally and externally.

The two participating organizations for this research were selected on the basis of comparability. First, both organizations are working in the field of international developmental, which can be defined as, engaging with economically disadvantaged regions in the world to empower people towards greater quality of life for humans and to address causes of poverty (University of Oxford, 2012). Second, both organizations currently have multiple operating chapters across Canada. Third, both organizations have comparable organizational capacities, budget, and membership.

Two different qualitative methods were utilized. First, the study examined the research questions at the organizational level. The data collected were comprised of publicly available annual reports, organizational documents, media releases, and official publications from the two participating INGOs released from 2005–2011. Any related documents published from third party sources were not included, as the project's intent was to capture the perspective from the organizations themselves.

In turn, content analysis was selected as the qualitative method used to analyze this data. Content analysis is a systematic process of classification coding, identifying themes or patterns from the identified codes, and creating an interpretation through the basis of these themes or patterns (Hsieh & Shannon, 2005). Content analysis is useful for this purpose as it can reveal either overt or subtle themes and characteristics of the text, which may be difficult to detect with casual observation (Neuman, 2004).

Second, qualitative interviews were conducted with directors, staff members, and volunteers to gauge their views, opinions, and experiences as it pertained to the research questions ($n = 10$). Qualitative interviewing provides inherent emphasis on depth, nuance, complexity, and roundedness in the data, as well as treating knowledge as situated and contextual (Mason, 2004). This helps the researcher in gaining a richer and more complex understanding of the participants' experiences and the social environment through those experiences (Keats, 2009). In organizational research, qualitative interviewing can facilitate an understanding of the perceptions, knowledge, and actions within an organizational setting which are often either directly and/or indirectly shaped through organizational structure, activities, processes, constraints, and existing social relations within (Campbell & Gregor, 2002).

The incorporation of these two qualitative methods within the research project further created the opportunity for triangulation. This is an important strategy often used as an effective means to offset the inherent biases in respective methods (Vitale, Armenakis & Field, 2008). Furthermore, these two methodologies offered two separate, yet interrelated levels of analyses.

The two sources of data collected were analyzed via NVivo 8. Analysis of the content was based on a careful and systematic observation driven by the rules denoted by the coding scheme. This study utilized a two-step deductive and inductive process of coding where nodes reflective of the research questions were created and any data that was relevant to any of the previously created nodes was coded. However, Miles and Huberman (1994) further explain that, as with all research projects, nodes often will change, develop and mold to the data that has been obtained. On the one hand, some nodes will not work and will end up decaying during the analysis process; on the other hand, other nodes will flourish and create the need to subcategorize existing nodes or develop new ones. Thus, a second inductive coding process was conducted where themes emerging from the data were coded into either newly created or supplemental nodes. Conceptual saturation was reached when no new categories could be generated from the data and existing codes have already been encapsulated into one or more of the created nodes (Kendall, 1999). This constituted the first iteration of the coding process.

The second iteration of the coding process involved clustering nodes together. The clustered nodes were then examined for their relationships to each other, elucidating the nature of those relationships. Finally, the findings were used to revisit the discussion within the literature to provide validation or corroboration to existing scholarship and to point out differences or gaps in current understandings of the phenomena (Kendall, 1999).

Due to the ethical considerations of the study, the identities of the participating organizations will be kept confidential in the findings of this study and the organizations will be distinguished by pseudonyms (“Organization A” and “Organization B”) and a general descriptive profile of the two participating organizations, their goals and the activities in which they engage.

Organization A is a development organization with religious origins, and its primary mandate is to support local NGOs working in the Global South, or “Partners” that promote and work toward alternatives to unfair social, political, and economic structures. Organization A is currently working in twenty-two countries in Africa, Asia, Latin America, and the Middle East. Organization A clusters their activities into five main categories: working with Partners to provide emergency humanitarian relief; developing and implementing longer term international developmental programs with Partners to improve the quality of life for those living in the Global South; building capacity and strengthening civil society in countries they work in; fundraising within Churches across Canada to support the organization and its Partners; and, launching campaigns to raise awareness, education, advocacy, and action in Canada.

Organization B is a fair trade organization, also with religious origins, with the aim to create economic opportunities for the most disadvantaged individuals, such as women and the disabled, in the Global South, as a strategy for poverty alleviation and development. In essence, fair trade constitutes an agreement between the trader and producer with the promise that the costs of production will be encompassed within the payment producers receive, enabling them to profit beyond covering the existing costs and providing fair payment to their labourers during the process. Organization B purchases traditional and cultural artistic products, made by producers or “Artisans” from South Asia, Southeast Asia, Africa, the Middle East, and Latin America and the Caribbean, and then markets and sells fair trade products via a network of branded retail stores across North America.

In addition to selling Fair Trade products, Organization B staff members and volunteers focus on providing education and awareness on how producers in the South are systemically disadvantaged within the

globalized free trade system and that fair trade presents an option to counteract that systemic impoverishment, in an effort to foster a culture whereby Fair Trade becomes increasingly part of the everyday culture in the North. Organization B also works with Southern umbrella NGOs that oversee the economic wellbeing of Artisans and provide them with holistic forms of support, such as childcare, medical services, training initiatives, and microloans to foster the process of developing independence and self-sufficiency.

FINDINGS

Descriptive and summative reporting of impact

The findings in this study largely reiterate the existing literature in that program evaluation and impact assessment are being actively conducted. In particular, results from programs and the changes coming as a result of programs implemented in the South were generally documented in both organizations' annual reports and media releases. For instance, in Organization A, the following illustrates the results of a project led by their Partner in Madagascar:

[The] ... project, which focused on increasing rice production in the Vatovavy Fitovinany region, has changed the lives of 1,000 rice-farming families. By being shown how to use carefully selected seeds and organic fertilizers, these families saw their production rise by 50 percent in 2009. As a result, they are no longer living in constant fear of hunger. (Organization A document, 2011)

Likewise, Organization B also documented the outcomes of programs implemented in areas of Punjab, Pakistan:

[This Partner] includes more than 820 families in 100 villages ... [sales] have enabled them to build and staff schools and to support a variety of village initiatives. There are now 1500 students in classes, over half of whom are girls. This is a marked achievement since the literacy rate for women in Pakistan is only 37% and substantially lower than that for women who live in villages. (Organization B document, 2011)

These two examples are representative of the type of reporting found in the annual reports, foundational documents, and media releases of both organizations. Based on the statistics and context provided within these reports, it can be deduced that some form of systematic data collection and assessment of impact on their initiatives exists within both organizations.

On the other hand, this type of reporting also raises questions of academic significance and rigour. As illustrated above, both organizations documented the results of their programs in a mostly short, descriptive, and anecdotal manner, which lacks the true breadth and depth of the knowledge accumulated. For instance, while both organizations presented statistics in their reports, questions of significance, whether confounding factors were taken into account in the analyses, and the relative effect size of their intervention were left unaddressed. Furthermore, these reports rarely presented the voices of Partners and of the local communities on the impact or lack of impact of the organizations' programs and interventions. Understandably, these thorough analyses would be more suitably placed in lengthier reports as opposed to annual reports and media releases, which are intended for smaller and quicker packets of information. However, these lengthier reports were absent for examination on both organizations' websites.

Program evaluation and impact assessment are being conducted primarily by head office rather than the entire organization.

At the ground level, it was found that directors, staff members, and volunteers of both organizations were all found to be familiar with the notion of conducting program evaluation and impact assessment. However, what was further uncovered was a gap in knowledge between those who worked in the national head offices of the organizations, and staff members and volunteers working at the local level.

Extensive knowledge was found amongst those working in the national head offices. In particular, one director described in detail their process of monitoring and assessment:

So, let's say ... in Colombia ... over the next five years, we help thirty-thousand families of Native people to settle down in their own land, in non-conflict areas of the country. And we implement that project together with [a Partner], every three months and every year (in a more formal way), those organizations report back to us what they have been doing and we can monitor whether or not the target...is being achieved or not. (Organization A director, February 28, 2011, Interview)

From this passage, it can be seen that the progress of work on the ground and information toward evaluation and assessment is actively collected and communicated between Organization A and its Partners. This director further explicated that:

[We] have hundreds of pages of results in different countries in terms of empowering the organization, in terms of bringing about significant change in the life...particularly of women and children in many countries, in terms of peace building and reconciliation in Afghanistan, in Iraq, in Colombia, in the Congo...many, many achievements. No, there is no doubt...depending on how much you want to spend on this and how much deeper you want to go. (Organization A director, February 28, 2011, Interview)

Once again, this statement reinforces the view that bountiful data, analyses, and results on their programs and their impact do exist within the organization.

However, staff members and volunteers were not active participants in the process of conducting program evaluation and impact assessment. When asked, one staff member from Organization A responded that, "Well, that's not [my] responsibility. That is definitely [part of national head office]...[that's] their job" (Organization A staff member, February 14, 2011 Interview). This passage suggests that staff members and volunteers working in the North do not appear to be involved in the process of evaluating or assessing their organization's activities carried out in the Global South; rather, this is the responsibility of a select few working within the national head office.

Even for activities conducted in the Global North, staff members and volunteers were only marginally involved with the process of evaluating or accessing their organization's work. One volunteer leader described their experience with data collection after a campaign:

After each campaign ... I receive a survey from national head office asking me to report back on the experience of the [awareness] campaign or the [fundraising] campaign, which is then sent back and compiled ... [the] difficulty with that, to be honest, is that sometimes I don't have all the information from all of my ... [volunteers] to get a really detailed report to

be able to send back to them, right? I'm not asking for my ... [volunteers] to report to me the same way that they're asking me to report to them, so I'm really basing on anecdotal evidence rather than getting a fairly detailed...so, what they get from me is ... kind of a general sense. (Organization A volunteer leader, March 11, 2011, Interview)

This passage suggests a disconnect between the need to assess the programs of Organization A, and the lack of communication and training provided to staff members and volunteers to be active participants in the evaluation and assessment process when it is expected. As the volunteer leader from Organization A articulated, there appears to be a lack of emphasis placed on the importance of evaluation and assessment amongst staff members and volunteers working on the ground. Consequently, volunteers likely have not been educated and trained to collect data during campaigns. Subsequently, this created the need to collect retrospective data, which is much less systematic and more likely to be anecdotal.

In addition to staff members and volunteers not being involved in the program evaluation and impact assessment process, it was further found that both organizations do not widely disseminate the results of their program evaluation and impact assessment back to their staff members and volunteers. When asked if it was common to receive updates from the Partners in the Global South or whether it would be up to the national head office to disseminate updates on the ground, one volunteer replied:

[It's] really focused through the [rollout of the] campaigns. It's really during those times that you would be getting those updates. Otherwise, you really do need to look for it on the web because they're all on there ... it's all on the website, what our partners are up to, but you kind of have to go digging for it. It comes to us from national office during the [campaign] season. (Organization A volunteer leader, March 11, 2011, Interview)

This suggests that other than the periods where staff members and volunteers execute a campaign focused on creating awareness and action in the North, there appears to be a general lack of awareness on the happenings on the ground amongst staff members and volunteers in Organization A, besides those actively engaged with their Partners within the national head office. The challenge may exist in the lack of mechanisms to disseminate knowledge throughout the organization and their consolidation within a segment of the organization.

The continuing challenge to quantify results

Finally, staff members and volunteers of both organizations articulated that a primary challenge to conducting evaluation and assessment within their respective organizations continues to be the emphasis placed on presenting the impact of their work as quantifiable outcomes. According to one volunteer, the challenge is that:

[It] is like an audit, right? But, in human development, you can't always do that ... [we're] not building so many schools. We're not building roads. So, it's hard to say because of the work we've done with this women's cooperative, this many women now can read or are now able to stand up for themselves and then educate their children or deal with spouses who are ... you know, there are all sorts of...you can't measure that and that's where we're really struggling because then again, more of your energy goes into how you're going to measure your results rather than actually how you're going to get results which you may not necessarily know at the beginning. (Organization A volunteer, June 23, 2011, Interview)

This common barrier is also echoed by staff and volunteers in Organization B, where one staff member commented that:

[Another NGO] can say, "We've build this many houses in Canada this year." And for most people love the fact that it's tangible, but it's a lot harder to say, "We've helped a woman who's left a battered situation, move into where she's now working and taking care of her family," because the measurements are not as obvious. They're must slower ... much longer process. (Organization B staff member, March 2011, Interview)

We suggest that both are eliciting parallel sentiments that the overreliance on quantitative measures leaves out the importance of the process of development work and the reactive and unpredictable elements that emerge. It is this type of unpredictable process that can lead to the more accurately articulated impact of development work. As the aforementioned Organization A volunteer describes, "The things where you may have started off here and gone this direction, and that it ends up going this way which, in the end, is wonderful, and it means so much more, but it's not what you said you were going to do. So, therefore, have you accomplished your goals?" (Organization A volunteer, June 23, 2011, Interview). This question is a crucial one for the work Organization B engages in. For instance, one commonly identified unintended impact of their work are Partners from diverse groups working together for the goals of fair trade in the Global South. One staff member describes this phenomenon of Christians, Hindus, and Muslims "working side-by-side and they were ... they were happy. They were laughing and talking. They knew ... I think, where their next meal was going to come from. They knew their children were in school" (Organization B staff member, April 6, 2011, Interview). Thus, the question remains: how would one organization properly measure that as an impact of their work?

Consequently, the concentration of reporting impacts using quantifiable measures remains an identified barrier for organizations when conducting evaluation and assessment. Specifically, the fixation on quantifiable measures leaves little room for using qualitative or formative data to report impact. This barrier becomes even more taxing when reporting impacts using quantifiable measures is a requirement for funding opportunities.

DISCUSSION

This study aimed to contribute to an increasingly pertinent discussion on the role of program evaluation and impact assessment within International Non-governmental Organizations (INGO) work. In turn, this study found that Organizations A and B are both actively conducting some form of program evaluation and impact assessment, which largely corroborates with the findings from previous studies (Campos et al., 2010; Morley, Vinson & Hatry, 2001; Travers, 2011). It was further found that members of both organizations were well aware of the need to demonstrate the effectiveness of INGO work and generally saw the use of program evaluation and impact assessment as being a positive tool to address this need. However, this study revealed that evaluation and assessment within these two organizations continued to be hindered with three challenges.

First, this study found that the results of the evaluation and assessment in Organizations A and B were reported through short descriptive and summative pieces in their organizational documents, annual reports, and media publications. The data, as presented in this manner, leave gaps in the knowledge communicated. More specifically, questions relating to academic rigour, such as methodology, significance, effect size, and replicability, are left unanswered. Understandably, the intended audiences of these documents are likely

members and supporters of the organization, as well as the general public; and the primary goal of these documents is to provide these audiences with succinct information on the activities and successes of the organization. Therefore, the addition of too much academic language may actually work to disengage the intended audience. However, this concern could potentially be overcome by the use of a two-pronged approach to reporting that includes an executive summary and a full report. Here, the executive summary can continue to serve the purpose of providing more general audiences with succinct information. Meanwhile, readers interested in the comprehensive evaluation and/or assessment framework can access the full report. By publishing and disseminating more sophisticated analyses, the knowledge base in the field of program evaluation and impact assessment within the nonprofit sector can continue to accumulate and strengthen. Furthermore, it can provide evidence to support INGOs' claims of impact and contribute to the overarching discussion on INGO effectiveness. As aforementioned, these reports may exist; however, they were not publicly accessible, which remains a limitation.

Alternatively, this finding may be an indication that there is a lack of expertise in conducting more sophisticated analyses within both organizations. As Travers (2011) previously found, the vast majority of INGOs in Canada used their own staff and/or interns and volunteers to conduct program evaluation and impact assessment. While both Travers (2011) and Dawson (2010) document that having internal staff members conduct evaluation and assessment, rather than external consultants, generates more autonomy, comfort, and trust within the INGO, the level of experience and expertise held by the staff member/intern/volunteer responsible for conducting evaluation and assessment remains unclear. Alaimo (2008) recommends that having a dedicated staff member with a background in program evaluation and impact assessment would be the ideal scenario; however, this is ultimately contingent upon whether organizations are able to successfully advocate for a position focused exclusively on evaluation and assessment to be part of the budget, which had been previously identified by the existing literature as a continuing challenge for INGOs.

Second, it was found that evaluation and assessment were primarily conducted within the national head offices of both Organization A and B, and that local staff members and volunteers were only marginally involved in the data collection process. In addition, it was revealed that the results of Organization A and B's evaluation and assessment of programs conducted in the Global South were not regularly communicated back to staff members and volunteers working in the North. When asked, a majority of staff members and volunteers working in the North stated that they were not involved in the dissemination process and thus did not feel they could confidently speak to questions pertaining to their perceptions of the impact of their organization's programs and interventions in the South.

One possible explanation for these findings is that the organizations' national head offices decided not to add to the already overburdened workload of local staff members and volunteers (the cases of both Organization A and B would support this). Among Organization A participants in the study, one local staff member was responsible for a province-wide jurisdiction, while the others were all volunteers who dedicated their time on top of their professional and personal responsibilities. Meanwhile, Organization B staff members consisted of store managers who worked between three-quarters to full-time in their role while the rest of the store was staffed by volunteers. In regards to conducting program evaluation and impact assessment, this raises concerns of time commitment and continuity for part-time volunteers, and too much additional workload for the small proportion of staff members in the local context.

On the other hand, improved efforts to educate staff members and volunteers in the evaluation and assessment process could be a possible solution. As underscored within the literature, program evaluation

and impact assessment are most effective when integrated into the organization's day-to-day operations and culture on the ground (Alaimo, 2008; Dawson, 2010). Dawson (2010) adds that staff members and volunteers need to be aware of purposes and procedures of the evaluation and assessment process in order for them to become active contributors to the process. We propose that a good starting point would be increased knowledge dissemination at the local level. By regularly communicating the findings and results of the organization's programs and interventions, INGOs begin to foster a culture of organizational learning based on the evidence collected on the ground (Edwards, 1997). Consequently, this can help staff members and volunteers begin the process of structured learning from experience, so that they are able to connect information and experience and turn it into knowledge and action (Edwards, 1997). As staff members and volunteers begin to understand and see the value of evaluation and assessment, they will be more likely to contribute to the process.

Finally, staff members and volunteers from both organizations echoed the difficulty in compartmentalizing the results of their work into finite and quantifiable indicators. Staff members and volunteers from both organizations echoed similar sentiments in articulating that their work in international development is not just "building this many number of houses" (Organization B volunteer, March 3, 2011, Interview). What is missing is the inclusion of reflective elements pertaining to change and meaningfully reporting on the process of change, which often does not operate in a linear and quantitative manner.

The pressure for this type of reporting can be largely attributed to the need for being accountable to donors and the guidelines set forth by the NPM model, which has been prominently cited within wider scholarship (Agg, 2006; Nicolau & Simaens, 2009; Smith, 2008; Szporluk, 2009). While completely changing the NPM paradigm does not appear likely in the near future, there are strategies INGOs can utilize to work within the existing context. Specifically, INGO leadership can inform their staff members and volunteers of the funding application process and the parameters of what is required to report in regards to the impact of their programs. Once again, through increased knowledge dissemination of the evaluation and assessment process within the organization, this opens the door for possibilities and opportunities for increased staff and volunteer assistance in this area of work within the organization. For instance, they can be involved in the planning and design of evaluation and assessment to better or more innovatively represent the impact of their work, while still adhering to the parameters of reporting. In addition, leadership, and staff members and volunteers, can begin to work together to find opportunities for incorporating evaluation and assessment into their organization's fabric, either in terms of day-to-day activities or a commitment to building capacity to conduct evaluation and assessment within the organization. Ultimately, this again leads to a shift in organizational culture.

Admittedly, by proposing increased staff member and volunteer involvement in adopting responsibilities related to conducting program evaluation and assessment, we reignite the issue of overburdened workloads. According to Dawson (2010), it is crucial to keep the evaluation and assessment process as simple and practical as possible to take up less staff time. We interpret this recommendation to convey the message that any program evaluation and/or impact assessment strategy must be designed in a manner that is simple to execute throughout the entire organization, while still effectively obtaining the necessary data in a systematic manner. A comprehensive organization-wide evaluation and assessment strategy, however, recrudesces the challenge of seeking expertise focused exclusively on program evaluation and impact assessment within INGO work. Without obtaining expertise in this specialized area, INGOs may continue to find themselves wrestling to find the right balance between conducting an ideal evaluation and assessment vis-à-vis a practical one.

CONCLUSION

In the twenty-first century, the discussion on the use of program evaluation and impact assessment to assist INGOs in documenting and demonstrating their effectiveness has emerged to become one of increasing importance in shaping the future of INGOs. This study found that while program evaluation and impact assessment are already being actively conducted within INGOs, there remains a gap between what has been recommended within the existing literature and the reality of those practices on the ground. Largely, the challenges faced by the two participating organizations in this study echo what has been documented in the existing scholarship. While we suggest that developing an organizational culture where program evaluation and assessment permeates all levels of the organization can increase understanding, buy-in, and participation amongst staff members and volunteers on the ground to better improve processes of data collection, the challenge of being able to design an ideal evaluation and assessment strategy will likely require INGOs to invest in some sort of expertise in the area of evaluation and assessment. Further exploring the value added by program evaluation and impact assessment specialists in the nonprofit sector may support INGOs in making the case for institutionalizing evaluation and assessment support in the future.

This study has limitations. Specifically, it must be acknowledged that it is not completely evident the degree to which the practices of the two INGOs studied are representative of other INGOs. This study would certainly also have benefited from the ability to interview staff members and / or volunteers from Partners working in the Global South. It was only during the data collection period that the realization of the knowledge gap of staff members in the North on the findings and results of evaluation and assessment conducted on activities and happenings in the South emerged. Future studies on gauging the perspectives of INGOs on the role of program evaluation and impact assessment within their work should focus on underscoring the voices of those working in the Global South and how their perspectives compare to those of staff members and volunteers working in the North.

ACKNOWLEDGEMENTS

The authors would like to thank the reviewers for their constructive and insightful feedback. We believe their input has resulted in a more balanced and focused article.

REFERENCES

- Abdel-Kader, M., & Wadongo, B. (2011). Performance management in NGOs: Evidence from Kenya. *Social Science Research Network*. URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1909863 [January 20, 2013].
- Agg, C. (2006). Winners or losers? NGOs in the current aid paradigm. *Development*, 49(2), 15–21.
- Alaimo, S. (2008). *Program Evaluation Capacity for Nonprofit Human Services Organizations: An Analysis of Determining Factors* (Doctoral dissertation). Bloomington, IN: Indiana University-Purdue University Indianapolis Scholar Works. URL: <https://scholarworks.iupui.edu/bitstream/handle/1805/1700/DISSERTATIONPDF9-19-08.pdf?sequence=1> [October 3, 2013].
- Atack, I. (1999). Four criteria of development NGO Legitimacy. *World Development*, 27(5), 855–864.

- Barber, M., & Bowie, C. (2008). How international NGOs could do less harm and more good. *Development in Practice*, 18(6), 748–754.
- Bouchard, M.J. (2009). The worth of social economy. In M. J. Bouchard (Ed.), *The worth of social economy: An international perspective* (pp. 19–34). Brussels: P.I.E. Peter Lang.
- Brown, L., & Moore, M. (2001). Accountability, strategy, and international nongovernmental organizations. *Nonprofit and Voluntary Sector Quarterly*, 30(3), 569–587.
- Campbell, D. (2002). Outcomes assessment and the paradox of nonprofit accountability. *Nonprofit Management and Leadership*, 12(3), 243–259.
- Campbell, M.L., & Gregor, F.M. (2002). *Mapping social relations: A primer in doing institutional ethnography*. Aurora, ON: Garamond Press.
- Campos, L., Andion, C., Serva, M., Rossetto, A., & Assumpção, J. (2010). Performance evaluation in Non-Governmental Organizations (NGOs): An analysis of evaluation models and their applications in Brazil. *Voluntas: International Journal of Voluntary and Nonprofit Organizations*, 22(2), 238–258.
- Dawson, A. (2010). A case study of impact measurement in a third sector umbrella organisation. *International Journal of Productivity and Performance Management*, 59(6), 519–533.
- Ebrahim, A. & Rangan, V. (2010). The limits of nonprofit impact : A contingency framework for measuring social performance (Harvard Business School General Management Unit Working Paper 10-099). *Social Science Research Network*. URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1611810 [January 20, 2013].
- Edwards, M. (1997). Organizational learning in non-governmental organizations: What have we learned? *Public Administration and Development*, 17(2), 235–250.
- Edwards, M., & Hulme, D. (1996). Introduction: NGO performance and accountability. In M. Edwards & D. Hulme (Eds.), *Beyond the magic bullet: NGO performance and accountability in the post-Cold War world* (pp. 1–21). West Hartford, CT: Kumarian Press.
- Fisher, W. (1997). Doing good? The politics and antipolitics of NGO practices. *Annual Review of Anthropology*, 26(1), 439–464.
- Fowler, A. (1996). Assessing NGO performance: Difficulties, dilemmas, and a way ahead. In M. Edwards & D. Hulme (Eds.), *Beyond the magic bullet: NGO performance and accountability in the post-Cold War world* (pp. 169–187). West Hartford, CT: Kumarian Press.
- Fowler, A. (1997). *Striking a balance: A guide to enhancing the effectiveness of non-governmental organizations in international development*. London, UK: Earthscan Publications Limited.
- Gibelman, M., & Gelman, S. (2004). A loss of credibility: Patterns of wrongdoing among nongovernmental organizations. *Voluntas: International Journal of Voluntary & Nonprofit Organizations*, 15(4), 355–381.
- Gosling, G., & Edwards, M. (2003). *Toolkits: A practical guide to planning, monitoring, evaluation and impact assessment*. London, UK: Save the Children.

- Harsh, M., Mbatia, P., & Shrum, W. (2010). Accountability and inaction: NGOs and resource lodging in development. *Development and Change*, 41(2), 253–278.
- Hsieh, H. & Shannon, S. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288.
- Keats, P. (2009). Multiple text analysis in narrative research: Visual, written, and spoken stories of experience. *Qualitative Research*, 9(2), 181–195.
- Kendall, J. (1999). Axial coding and the grounded theory controversy. *Western Journal of Nursing Research*, 21(6), 743–757.
- Kilby, P. (2004). Is empowerment possible under a new public management environment? Some lessons from India. *International Public Management Journal*, 7(2), 207-225.
- Lecy, J., Schmitz, H., & Swedlund, H. (2011). Non-governmental and not-for-profit organizational effectiveness: A modern synthesis. *Voluntas: International Journal of Voluntary & Nonprofit Organizations*, 23(2), 434–457.
- Lockie, S. (2001). SIA in review: setting the agenda for impact assessment in the 21st century. *Impact Assessment and Project Appraisal*, 19(4), 277–287.
- Mason, J. (2004). *Qualitative Researching* (2nd ed.). London, UK: Sage Publications.
- Miles, M. & Huberman, A. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Morley, E., Vinson, E., & Hatry, H. (2001). *Outcome measurement in nonprofit organizations: Current practices and recommendations*. Washington, DC: Independent Sector. URL: http://www.iyi.org/resources/mentor/pdf/Outcome_Measurement.pdf .
- Moxham, C. (2009). Performance measurement: Examining the applicability of the existing body of knowledge to nonprofit organisations. *International Journal of Operations & Production Management*, 29(7), 740–763.
- Neuman, W. (2004). *Basics of social research: Qualitative and quantitative approaches*. Boston, MA: Pearson Education, Inc.
- Nicolau, I., & Simaens, A. (2009). The evaluation of social solidarity organizations in the Portuguese context. In M. J. Bouchard (Ed.), *The worth of social economy: An international perspective* (pp. 193–212). Brussels: P.I.E. Peter Lang.
- Parks, T. (2008). The rise and fall of donor funding for advocacy NGOs: Understanding the impact. *Development in Practice*, 18(2), 213–222.
- Perret, B. (2009). Evaluating the social economy. Clarifying complex rationality. In M. J. Bouchard (Ed.), *The worth of social economy: An international perspective* (pp. 35–42). Brussels: P.I.E. Peter Lang.

Yu & McLaughlin (2013)

Sarker, A. (2005). New public management, service provision and non-governmental organizations in Bangladesh. *Public Organization Review: A Global Journal*, 5(3), 249–271.

Serva, M., Andion, C., Campos, L., & Onozato, E. (2009). Evaluation of the social economy in Brazil: An analysis of the practices in some NGOs. In M. J. Bouchard (Ed.), *The worth of social economy: An international perspective* (pp. 171–192). Brussels: P.I.E. Peter Lang.

Smith, M. (2008). International non-governmental development organizations and their Northern constituencies: Development education, dialogue and democracy. *Journal of Global Ethics*, 4(1), 5–18.

Spar, D., & Dail, J. (2002). Of measurement and mission : Accounting for performance in non-governmental organizations. *Chicago Journal of International Law*, 3(1), 171–181.

Szporluk, M. (2009). A framework for understanding accountability of international NGOs and global good governance. *Indiana Journal of Global Legal Studies*, 16(1), 339–361.

Travers, S. (2011). *Canadian civil society organizations influencing policy and practice: The role of research*. Ottawa, ON: International Development Research Centre. URL: <http://www.idrc.ca/EN/Documents/Stacie-Traver-working-paper-5-final.pdf> .

University of Oxford (2012). *International Development*. URL: <http://www.careers.ox.ac.uk/options-and-occupations/sectors-and-occupations/international-development/> [October 3, 2013].

Vitale, D., Armenakis, A., & Field, H. (2008). Integrating qualitative and quantitative methods for organizational diagnosis: Possible priming effects? *Journal of Mixed Methods Research*, 2(1), 87–105.

Walsh, E. & Lenihan, H. (2006). Accountability and effectiveness of NGOs: Adapting business tools successfully. *Development in Practice*, 16(5), 412–424.

ABOUT THE AUTHORS / LES AUTRES

Stan Yu is Research and Program Evaluation Analyst for the Gwenna Moss Centre for Teaching Effectiveness and teaches in the Department of Sociology, University of Saskatchewan. Email: stan.yu@usask.ca .

Darrell Arthur McLaughlin is Associate Professor of Sociology at St. Thomas More College (STM), University of Saskatchewan. Email: dmclaughlin@stmcollege.ca .