Intersections and Innovations

Change for Canada's Voluntary and Nonprofit Sector

The Muttart Foundation

ISBN: 978-1-897282-30-4

The Muttart Foundation

Acknowledgements

For far too long, Canada has lacked a comprehensive resource examining Canada's charitable sector. That has now ended.

The Muttart Foundation has spent many years focusing on building the capacity of charities in this country. The publication of this collection is another contribution to that effort. By understanding more about itself, the sector can continue to develop and find new ways to serve Canadians and those in need outside our nation.

The authors of these essays bring different perspectives on the role and inner workings of Canada's charities. Collectively, they bring an unprecedented insight into the work of organizations whose diversity is exceeded only by their desire to serve.

It is difficult to express adequate appreciation to Dr. Susan Phillips of Carleton University for her leadership of this project. She has been a source of encouragement, persuasion, cajoling and improving authors from across the country. Her efforts now bear fruit as we make this material available to students, academics, practitioners and others interested in the history and future of Canada's charities.

Amanda Mayer of the Lawson Foundation volunteered at the outset to be the administrative overlord of the project, keeping the editors and authors up to date and keeping track of various versions of articles. We are so grateful for her skills, her patience and her friendship.

None of this would have been possible, of course, without the work of authors, themselves academics and/or practitioners. They took time from their schedules to contribute to a resource we hope many will find valuable.

Lesley Fraser did an incredible job in editing the various chapters and ensuring consistency. And Don Myhre of P40 Communications has again brought his talent to the fore in providing an attractive design for a Muttart publication.

The work of all of these individuals has come together in this resource which we dedicate to all of those in, or interested in, Canada's charitable sector.

Malcolm Burrows, President

Bob Wyatt, Executive Director

The Muttart Foundation

The Muttart Foundation

This book may be cited as:

Phillips, Susan D. and Wyatt, Bob (Eds) (2021) *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector.* Edmonton, AB, Canada: Muttart Foundation

Individual Chapter Citations

Chapter 1

Phillips, Susan D. and Wyatt, Bob (2021) Intersections and Innovations: Change in Canada's Voluntary and Nonprofit Sector. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector.* Edmonton, AB, Canada: Muttart Foundation

Chapter 2

Lasby, David and Barr, Cathy (2021) State of the Sector and Public Opinion about the Sector. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector.* Edmonton, AB, Canada: Muttart Foundation

Chapter 3

Marshall, Dominique (2021) Four Keys to Make Sense of Traditions in the Nonprofit Sector in Canada: Historical Contexts. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector.* Edmonton, AB, Canada: Muttart Foundation

Chapter 4

Wyatt, Bob (2021) It Should Have Been So Simple: The Regulation of Charities in Canada. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector.* Edmonton, AB, Canada: Muttart Foundation

Chapter 5

Chan, Kathryn and Vander Vies, Josh (2021) The Evolution of the Legal Meaning of Charity in Canada: Trends and Challenges. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector.* Edmonton, AB, Canada: Muttart Foundation

Chapter 6

Manwaring, Susan and Kairys, Katrina (2021) Regulating Business Activity. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector.* Edmonton, AB, Canada: Muttart Foundation

Chapter 7

Phillips, Susan D., Dougherty, Christopher, and Barr, Cathy (2021) The Fine Balance of Nonprofit Sector Self-Regulation: Assessing Canada's Standards Program. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector.* Edmonton, AB, Canada: Muttart Foundation

Chapter 8

Charters, Owen (2021) Board Governance in Practice. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector.* Edmonton, AB, Canada: Muttart Foundation

Chapter 9

Grasse, Nathan and Lam, Marcus (2021) Financing Canadian Charities: The Conditional Benefits of Revenue Diversification. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector.* Edmonton, AB, Canada: Muttart Foundation

Chapter 10

Hale, Sharilyn (2021) Giving and Fundraising Trends. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector.* Edmonton, AB, Canada: Muttart Foundation

Chapter 11

Glogovac, Marina (2021) New Technologies and Fundraising. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector.* Edmonton, AB, Canada: Muttart Foundation

Chapter 12

Fontan, Jean-Marc and Pearson, Hilary (2021) Philanthropy in Canada: The Role and Impact of Private Foundations. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector.* Edmonton, AB, Canada: Muttart Foundation

Chapter 13

Khovrenkov, Iryna (2021) Canada's United Way Centraide as a Community Impact Funder: A Reinvention or a Failed Endeavour? In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector.* Edmonton, AB, Canada: Muttart Foundation

Chapter 14

Harji, Karim and Hebb, Tessa (2021) Impact Investing in Canada: Notes from the Field. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector*. Edmonton, AB, Canada: Muttart Foundation

Chapter 15

Raggo, Paloma (2021) Leadership in the Charitable Sector: A Canadian Approach? In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector*. Edmonton, AB, Canada: Muttart Foundation

Chapter 16

Fredette, Christopher (2021) Planning for Succession in the Interests of Leadership Diversity: An Avenue for Enhancing Organizational Diversity, Inclusion, and Equity. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector*. Edmonton, AB, Canada: Muttart Foundation

Chapter 17

Akingbola, Kunle and Toupin, Lynne (2021) Human Resource Management in the Canadian Nonprofit Sector. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector*. Edmonton, AB, Canada: Muttart Foundation

Chapter 18

Uppal, Pamela and Febria, Monina (2021) Decent Work in the Nonprofit Sector. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector*. Edmonton, AB, Canada: Muttart Foundation

Chapter 19

Thériault, Luc and Vaillancourt, Yves (2021) Working Conditions in the Nonprofit Sector and Paths to Improvement. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector*. Edmonton, AB, Canada: Muttart Foundation

Chapter 20

Russell, Allison, Speevak, Paula, and Handy, Femida (2021) Volunteering: Global Trends in a Canadian Context. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector*. Edmonton, AB, Canada: Muttart Foundation

Chapter 21

Shier, Micheal L. (2021) Social Innovation and the Nonprofit and Voluntary Sector in Canada. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector*. Edmonton, AB, Canada: Muttart Foundation

Chapter 22

McCort, Kevin and Phillips, Susan D. (2021) Community Foundations in Canada: Survive, or Thrive? (with apologies to lawn bowlers). In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector.* Edmonton, AB, Canada: Muttart Foundation

Chapter 23

Murphy, Colette (2021) Community Wealth Building: A Canadian Philanthropist's Perspective. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector.* Edmonton, AB, Canada: Muttart Foundation

Chapter 24

Doberstein, Carey (2021) Collaboration: When to Do It and How to Do It Right. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector.* Edmonton, AB, Canada: Muttart Foundation

Chapter 25

Munshi, Shereen and Levi, Elisa (2021) Indigenous Peoples, Communities, and the Canadian Charitable Sector. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector.* Edmonton, AB, Canada: Muttart Foundation

Chapter 26

Stauch, James, Glover, Cathy, and Stevens, Kelli (2021) The Business–Community Interface: From "Giving Back" to "Sharing Value." In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector.* Edmonton, AB, Canada: Muttart Foundation

Chapter 27

Laforest, Rachel (2021) Transforming Health and Social Services Delivery Systems in Canada: Implications for Government–Nonprofit Relations. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector.* Edmonton, AB, Canada: Muttart Foundation

Chapter 28

White, Deena (2021) Contentious Collaboration: Third Sector Service Delivery in Quebec. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector.* Edmonton, AB, Canada: Muttart Foundation

Chapter 29

Levasseur, Karine (2021) Policy Capacity: Building the Bricks and Mortar for Voluntary Sector Involvement in the Public Policy Process. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector.* Edmonton, AB, Canada: Muttart Foundation

Chapter 30

Houston, Sandy (2021) Evolving Relationships with Government: Building Policy Capacity. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector*. Edmonton, AB, Canada: Muttart Foundation

Chapter 31

Northcott, Allan (2021) Reflections on Teaching Public Policy Advocacy Skills. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector*. Edmonton, AB, Canada: Muttart Foundation

Chapter 32

Lauzière, Marcel (2021) A Lever for Change: How Foundations Can Support Public Policy Advocacy. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector*. Edmonton, AB, Canada: Muttart Foundation

Chapter 33

Ruff, Kate (2021) Social and Environmental Impact Measurement. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector*. Edmonton, AB, Canada: Muttart Foundation

Chapter 34

Lenczer, Michael, Bourns, Jesse, and Lauriault, Tracey (2021) Big Data Won't Save Us: Fixing the Impact Evaluation Crisis. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector*. Edmonton, AB, Canada: Muttart Foundation

Chapter 35

Herriman, Margaret (2021) Social Media and Charities in Canada. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector*. Edmonton, AB, Canada: Muttart Foundation

Chapter 36

Riseboro, Caroline (2021) The Overhead Myth: The Limitation of Using Overheads as a Measure of Charity Performance. In Susan D. Phillips and Bob Wyatt (Eds.), *Intersections and Innovations: Change for Canada's Voluntary and Nonprofit Sector*. Edmonton, AB, Canada: Muttart Foundation

Part III Innovation and Intersections

Community and Corporate Intersections

Intersections with Governments: Services and Policy Engagement

Measuring Impact and Communicating Success

The Muttart Foundation

Measuring Impact and Communicating Success

Chapter 33

Social and Environmental Impact Measurement

Kate Ruff
Carleton University

This chapter discusses best practices and key challenges of performance assessment in charities, nonprofits, and social enterprises in Canada, with a focus on the practical steps organizations can take to make performance assessment work better for them.

The main argument is that performance assessment can be relatively easy when done for a specific purpose at a single organization or project but that it becomes much more challenging when measures must fulfil multiple purposes and must exist within an ecosystem of interconnected organizations. Given that most organizations (dare I say all?) exist in this more challenging scenario, good performance assessment requires an understanding of the tensions that arise when measures are used for multiple purposes and multiple organizations.

The chapter begins with a description of impact measurement in its simplest form: when undertaken for a single purpose and a single organization. As noted above, impact measurement in this scenario is relatively easy. It is easy in the way that running a marathon is "easy." It can take hard work, but there is no mystery. This is something people know how to do, and many have done before. If done frequently, it really does become easy.

The second section introduces complexity into the simple scenario. It focuses on the multiple roles that performance measurement is called upon to fulfill: learning, impact management, accountability, and communication. I identify the conditions under which performance measurement can fulfill those roles and highlight the perverse consequences that arise when a measurement system designed for one purpose is called upon to fulfill another. The key message is that measurement is not a panacea, nor a crystal ball. Best practice suggests a diverse range of metrics and a lot of human discretion in their interpretation.

The third section introduces even more complexity. It examines measurement within an ecosystem of interconnected organizations. Impact measurements flow among organizations to facilitate communication, coordination, and collaboration. It is helpful if the organizations are speaking the same language – if they use similar measures. What are the consequences of uniform or standardized metrics? I argue that a balance must be struck between relevance, which requires flexibility, and uniformity.

Many of the performance measurement tools created for performance (or impact) measurement assume the easy scenario: a specific purpose and single organization. One reason that these tools often fall short of expectations is that organizations do not, and cannot, operate in such a simplified world. Organizations are interconnected. Measures are used for multiple purposes. This reality makes measurement more challenging than existing resources acknowledge. The chapter concludes with some implications for Canadian nonprofits, expressed as seven tips for getting the most out of performance assessment.

Impact Measurement Basics

Performance assessment is easy when undertaken by a single organization for a single purpose. This section presents the causal model and measurement process that underpin most measurement methods. It is not a how-to guide, but links to further resources are provided.

Defining "Impact"

Instead of defining impact, I think it preferable to be aware of different ways that people use the term "social impact." The most common meanings of social impact are:

- **Impact as a generic word for any kind of social or environmental result.** Socially responsible companies, social enterprises, and impact investors use "social impact" to refer to social or environmental outputs and outcomes and "impact measurement" to refer to the tracking of those items (e.g. Mudaliar, Pineiro, Bass, & Dithrich, 2017; GRI, 2016). In this definition, impact measurement is synonymous with what nonprofits might call performance measurement. Many nonprofits have followed suit such that this generic use is now the most common meaning of impact, although there is a push to connect impact to contribution.
- **Impact as contribution.** Impact is used to describe change *as a result* of an organization's activities, meaning the change would not have happened anyway. Impact measurement in this context refers to methods that allow claims of attribution (White, 2010) by making use of control groups and experimental design. In this definition, impact measurement is different from outcome measurement because it can identify contribution. This usage is particularly common among evaluators (Hansen, Klejnstrup, & Andersen, 2013) and was recently embraced by the Impact Management Project (IMP), a global forum that seeks to build a consensus about measurement of impact.
- **Impact as significant change.** The term "impact" is used to describe the follow-on (Campbell, 2002; White, 2010) or long-term fundamental (W.K. Kellogg Foundation,

2004) effects of an organization's work, such as effects on need and populations (Poole et al., 2000). In this definition, impact measurement refers to the effects that come after outcomes – often, the long-lasting effects.

- **Impact as valuation.** Impact can refer to the outcomes expressed as financial values. For example, the "social return on investment" (SROI) method calculates impact by multiplying outcomes with financial proxies and estimates of attribution (SROI, 2012). This definition of impact incorporates both contribution and relative importance (value) of different outcomes to stakeholders. Impact refers to the outcomes that matter, or better: outcomes weighted by how much they matter to stakeholders.

A Common Set of Foundational Practices

Almost all impact-measurement approaches recommend the same basic practices, as shown in Figure 1 (Arvidson & Lyon, 2014; also see Likert & Maas, 2015). The steps are not as sequential as implied in this figure. For example, data collection is ongoing throughout the cycle. Engaging stakeholders is integral to every practice.

- **Engage stakeholders:** Stakeholders are those who affect and are affected by your organization (Mitchell, Wood, & Agle, 1997). Primary stakeholders may include members, clients, beneficiaries, staff, funders, and subsets within these groups. Identifying and maintaining an open dialogue with stakeholders is widely recognized as a best practice (Likert & Maas, 2015). Engaging with stakeholders throughout the measurement cycle can help organizations to more *accurately* understand what changed (Mook, Quarter, & Richmond, 2003), and it can be an effective way of creating change through *sense-making* (Baker & Schaltegger, 2015). Communicating results with all stakeholders (not just funders) allows them *to hold organizations to account* (Chen, 2012).
- **Plan your change:** Best practice suggests that – in dialogue with stakeholders – the organization should determine what it intends to achieve and how it will get there (Likert & Maas, 2015). For those who define effectiveness as *responsiveness to stakeholders* (see Sowa, Seldon, & Sandfort, 2004: 713), objectives are emergent, fluid, and change often. For those who define effectiveness as *goal attainment* (see Sowa, Seldon, & Sandfort, 2004: 713; Liket & Maas, 2015: 271), objectives should be stable.
- **Use performance measures:** Express objectives and planned activities in terms of outcomes and outputs, and identify indicators to signal progress. It is best practice to consider both intended and unintended effects to get a broad sense of *what changed* (Campbell, 2002; White, 2010). *Indicators are verifiable measures* that mark or signal that an output, outcome, or impact has been achieved (Melnyk, Steward, & Swink, 2004). Indicators can be quantitative or qualitative. Typically, they do not depict the outcome or impact in its entirety but signal that it has likely happened to some degree.
- **Collect useful data:** Strive to make data collection routine. Ensure that what data you do collect is stored in useful ways. It is time-consuming to collate data from notebooks and Word files on individual computers. Instead, opt for shared spreadsheets, online forms, and specialized software like Sametrica and Impact Dashboard. Periodically review the data to make sure that everything collected is useful in some way.
- **Assess and gauge impact:** Ask if impact is being achieved as expected and if it is

valuable to stakeholders. Indicators never tell the whole story. Take time to make sense of the data in context, with risks and mitigating factors. Even if the results are not measured using experimental or quasi-experimental design (White, 2010), contemplate how much of them can be attributed to the organization (Mayne, 2001). If using the SROI method, it is at this stage that outcomes will be valued with financial proxies (SROI, 2012). For other methods, it is at this stage that organization and project managers draw summative conclusions from that data.

- **Report:** It is best practice to share assessments with stakeholders and to adapt programs and priorities for ongoing impact management (Liket & Maas, 2015). Responding doesn't follow automatically from performance measurement, however. Leaders must decide what to report, what can be learned, and how to adapt. Reports are important to close the loop. With reports, stakeholders can better participate in program design (Benjamin, 2013). Research suggests that nonprofits could do a better job in public reporting (Saxton & Guo, 2011; Dhanani & Connolly, 2012; Connolly & Hyndman, 2013).

Figure 1: The Foundational Practices

These five practices are common to most impact-measurement tools.

(Source: adapted from [Ontario Impact Measurement Task Force Action Plan)]

Resources: There are numerous resources to help organizations start on their performance assessment journey with the common process:

- Common Foundations (above)
- Demonstrating Value Workbook
- SROI Guide
- Territoires innovants en économie sociale et solidaire
- Universal Standards for Social Performance Management: implementation guide

The Causal Model

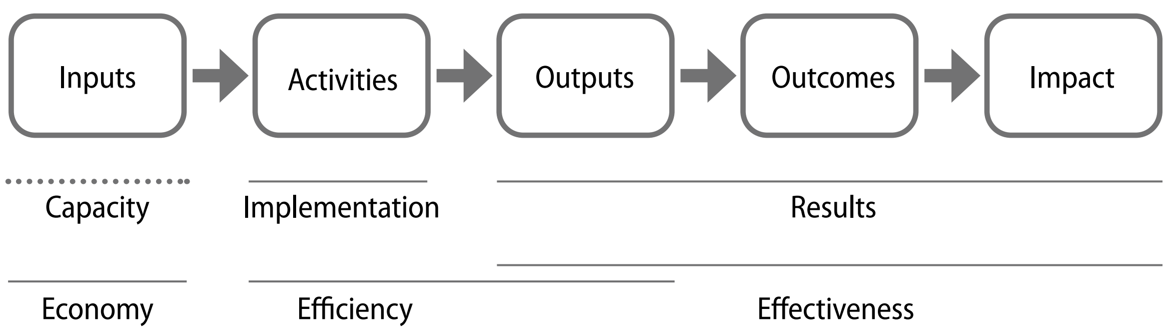
The most foundational practice, "plan your change," recommends that performance assessment be organized around a causal model, one in which attribution is possible. The model depicts how an organization's work will lead to the change it seeks (Weiss, 1997). Such models are at the heart of almost every measurement tool available, often under several variations, including the theory of change (Weiss, 1997), outcome mapping (e.g. Earl, Carden, & Smutylo, 2001), chain of results (e.g. Wimbush, Montague, & Mulherin, 2012; Mayne, 2001), impact map (SROI, 2012), and the logic model (e.g. W.K. Kellogg Foundation, 2004).

These various versions share a common vocabulary.

- **Inputs** are resources mobilized to effect change. Examples include staff, donations of money and goods, investments, volunteer labour, and community support. The list of inputs should be comprehensive and appropriate for the type of program (Poole, Nelson, Carnahan, Chepenik, & Tubiak, 2000).
- **Activities** *describe* what the organization does or will do, such as teach classes or provide counselling. Activities should have a logical link to outputs and be sufficient to achieve outcomes (Poole et al, 2000).
- **Outputs** are the direct products of program activities; they measure the type, level, and quantity of delivery (W.K. Kellogg Foundation, 2004; McLaughlin & Jordan, 1999).
- **Outcomes** are the benefits or results of activities. Outcomes can be short-term or longer-term. They should focus on what changed (Poole et al., 2000) and specify *what* will change, for *whom*, and by *how much*.
- **Impact,** as discussed above, has many meanings. Here we use it to mean "positive and negative, primary and secondary long-term effects produced by an intervention, directly or indirectly, intended or unintended" (IMP Glossary; White, 2010).

Figure 2: The Logic Model

A causal model describes how an organization's work will lead to the change it seeks:



(Source: Adapted from W.K. Kellogg Foundation [2004: 3] and Brignall & Modell [2000: 286])

Causal models differ slightly in their focus. More recent variations – for example outcome mapping and chain of results – emphasize outcomes and impacts. This reflects a prevailing discourse that organizations should move away from tracking outputs and place greater attention on *impact*, however nebulously defined.

The academic literature has a more holistic understanding of *effectiveness*, however. Scholars have noted that capacity and implementation are crucial to creating and sustaining effectiveness (Sowa, Seldon, & Sandfort, 2004; Herman & Renz, 1999). Research suggests that organizations that track capacity (inputs), the particulars of how work is implemented (activities), and results (outputs, outcomes, and impact) are better equipped to manage and adapt and thereby achieve results (Keevers, Treleaven, Sykes, & Darcy, 2012). Models that focus exclusively on outcomes and impacts risk overlooking how much management effort is spent securing the inputs and designing activities, and thus mask the effects of input-related successes (e.g. grants and hiring) on subsequent outcomes and impact success.

For this reason, I recommend that nonprofits start with the logic model (Figure 2). It includes all aspects of the organization's work from inputs, activities, and outputs through to impact (however defined). It can be used to track capacity, implementation, economy, and efficiency, which together enrich a nonprofit manager's understanding of how best to achieve results. Once the logic model is well understood, outcome- and impact-focused models can be layered on top of the logic model to give managers greater insight into the right-hand side of the process.

Resources for creating a causal model:

Creating Your Theory of Change: NPC's Practical Guide

Demonstrating Value Impact Mapping Worksheet

Territoires innovants en économie sociale et solidaire

The Roles of Social Impact Measurement

The above section outlined a straightforward approach to impact measurement. Most practice-oriented impact guides focus on this approach – impact measurement for a single purpose within a single organization. The problem with this simple scenario is that it doesn't reflect the reality nonprofits operate in. In real life, measures must serve multiple purposes, and in real life measures interface across an ecosystem of interconnected organizations. In this section, I examine the limitations, tensions, and dysfunctions that arise when measures are used for multiple purposes within a single organization.

Organizations do performance assessment for several reasons. A classic characterization of measurement is to prove and improve. Two slightly different reasons, as noted by Mayne (2001), are understanding and reporting. Melnyk, Stewart, & Swink (2004) chose three: control, communication, and improvement. (Note that they are using "control" in the accounting sense, meaning to ensure that things go as planned, such as a certain number of outputs within a specific budget. I will refer to this as "manage.") In addition to these *instrumental* roles, measures can be used in an *expressive* role to communicate the values and beliefs of the organization and those within it (Chenhall, Hall, & Smith, 2017). In this sense, working through measurement aims to bring people into agreement (Chenhall, Hall, & Smith, 2013).

Taken together, then, scholarly literature identifies these four slightly overlapping reasons for measurement:

- **Learn:** Get better at having an impact and share knowledge to help others improve.
- **Manage:** Track budgets, activities, and outputs, and link these to results.
- **Account:** Inform stakeholders so they too can learn, manage, account, and communicate.
- **Communicate:** Externally, tell a story to community members and policy-makers to mobilize resources and support; internally, develop a common understanding within the organization about priorities and goals.

What Measurement Can Reasonably Do

How effective can measurement be at helping organizations to learn, manage, account, and communicate? Simply stated, measurement can provide clear answers when there is high certainty of both causal relationships and objectives of the program (Burchell, Clubb, Hopwood, Hughes, & Nahapiet, 1980; Ebrahim & Rangan, 2010; see Figure 3).

- *Certainty of cause and effect* refers to the relationship between an organization's actions and results. Outputs and immediate outcomes are often highly certain (Ebrahim & Rangan, 2010). For example, feeding people reduces their hunger. Further down the causal chain, the relationship between actions and results becomes less certain.
- *Certainty of objectives* refers to consensus on goals (Ebrahim & Rangan, 2010). It is common to have consensus within a single organization. Disagreement over objectives is common across organizations. In the case of food, some people may disagree that the objective is to alleviate immediate hunger, instead aiming for improved nutrition or changing the socioeconomic conditions that create hunger in the first place.

Figure 3: Decision-Making and Organizational Uncertainty

		Certainty of cause and effect Will these actions achieve the aims?	
		Low certainty	High certainty
Certainty of objectives Is this the right thing to aim for?	High certainty	Decision by judgment Measures help organizations LEARN	Decision by computation Measurement can provide ANSWERS
	Low certainty	Decision by inspiration Measures help organizations RATIONALIZE	Decision by compromise Measures help organizations ADVOCATE

(Source: Adapted from Burchell, Clubb, Hopwood, Hughes, & Nahapiet, 1980)

Measurement can provide answers only when objectives are agreed upon and cause and effect are well understood (Burchell et al., 1980). This is often the case when an organization is making decisions about outputs.

If there is some uncertainty, measures can only contribute information to a process of analysis.

- When there is uncertainty about objectives, measurement can inform (but not replace) judgment. In these situations, impact measurement can contribute to learning (Burchell et al., 1980) and, as more is known about cause and effect, it may *eventually* help answer questions.
- When there is uncertainty about cause and effect, measurement can inform (but not replace) negotiation and compromise. In these situations, measures are used like ammunition (Burchell et al., 1980), with each party amassing an arsenal and firing off facts. The measures don't so much answer a question as empower a position.
- When there is uncertainty of objectives *and* cause and effect are not known, measurement is of limited use. At most, measures are mobilized after the fact to make an intuitive process *look* scientific and to rationalize decisions already made (Burchell et al., 1980).

Notice that accountability, communication, and management are not mentioned in Figure 3. How effective measurement is for those tasks will depend on the level of certainty. Some accountability and communication documents will use measurement as after-the-fact rationalizations; others will use measures as ammunition to persuade funders and policy-makers that the project was a success and funding should be renewed.

Some measurement proponents imagine a world where all decisions can be resolved by computation. When existing measures fail to provide unambiguous answers, additional measures are added. When answers remain murky, measures are changed and more added. And on it goes. This is the expectations gap. It is useful to keep in mind that there are many questions that data will never answer. Only when objectives and cause and effect are clear can measures provide answers.

When Measurement Plays Many Roles

Numerous tensions and dysfunctions arise when measurement is called upon to play many roles in an organization, which it almost always is. Four tensions and dysfunctions frequently occur:

- **Dynamic tension:** Effective measurement must be both stable and changing, creating a dynamic tension. For accountability and performance management, measures need to be stable. This is because comparisons over time require that measures stay the same over time. Improvement and innovation, however, require measures to be flexible (Johnston, Brignall, & Fitzgerald, 2002; Melnyk, Stewart, & Swink, 2004). Measures need to change if the program objectives change, and might need to change if activities change. As soon as measures change (good for learning and innovation), the accountability and performance-tracking systems break.
- **Complexity tension:** Measures need to be both complex and simple. Any given set of indicators necessarily excludes some things and prioritizes others. This is helpful when it gives focus to organizations and clarifies roles for employees (Hall, 2002). It can be incredibly harmful, though, when the same types of things are consistently excluded. A larger set of measures will provide a more holistic view (better for communication, learning, and accountability), but too much measurement distracts from the real work (Melnyk, Stewart, & Swink, 2004). For example, Keevers and colleagues (2012) observed how an emphasis on measuring the outcomes (health, security, etc.) of homeless youth undermined the ability of workers to build the relationships with the youth that were necessary to help the youth feel secure, supported, and healthy. Almost every set of impact measures is too simple in some ways and too complex in others.
- **Accountability versus learning:** For effective learning, managers need to be able to look at the data and engage in honest reflection on how to make things better (Ebrahim, 2005). When performance measures drive bonuses to program managers or influence decisions to increase a program's budget, there are incentives to pick measures that managers can control (Mayne, 2001; Theuvsen, 2004) and that are easiest to perform well on (Herman & Renz, 1999). Thus, when measures are part of accountability systems, managers focus on achieving rather than reflecting and improving (Phillips & Karlan, 2018), transforming what could be a holistic learning system into a hierarchical accountability system (O'Dwyer & Underman, 2008). The tension is created because one person's learning system is another person's accountability system: the executive director learns about her organization based on data that her program managers submit (accountability). Her funders learn about their work based in part on the data she submits. Learning and accountability are not either/or, but intertwined.
- **Metric corruption:** One of the greatest challenges of measurement is that the more important a metric, the less effective it becomes (DiMaggio, 2001; Campbell, 1979). DiMaggio (2001: 259) calls this the "flamingo problem" because, like the mallets in *Alice in Wonderland*, "indicators often develop lives of their own, and the meanings change as they are measured." In a statement now referred to as Campbell's Law, "The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor" (Campbell, 1979: 85). The

greater the consequences attached to the measure, the more the measure becomes the goal in itself. Like "teaching to the test," organizations seek to increase performance against the indicator rather than seek to influence the phenomenon that the indicator is meant to measure. Slowly the indicator becomes less connected to the phenomenon. Test scores no longer signify a broader notion of learning. Useful indicators often become problematic indicators by virtue of having once been useful.

The above tensions cannot be avoided. There is no known solution. There are, however, strategies that can mitigate their effects.

Strategies for Performance Measurement That Balance the Many Purposes

To minimize harm, best practice recommends a mix of measures across the following categories (Hall, 2008; Herman & Renz, 1999):

- Select at least one outcome indicator for each area or topic that matters (Brignall & Modell, 2000; Melnyk, Steward, & Swink, 2004). Some performance-measurement frameworks, such as the sustainable livelihoods approach, suggest topic areas; otherwise, the outcomes in the organization's causal model are the topics.
- Choose some "lead indicators" that give advance insight into how things are going, as well as indicators that show how things went. This can be thought of as the indicator's *tense* (Melnyk, Steward, & Swink, 2004). For example, if job placement is important, consider tracking both the number of interviews and the number of placements.
- Include output or near-term outcomes that capture the quality of the work. For many organizations, *how* they go about their work affects the results they are likely to achieve (Keevers et al., 2012). Consider, for example, whether staff feel they have enough time to establish rapport with clients.

In combination with a range of indicators, culture and leadership can help to keep an organization focused on forward-looking action. For this, a "good enough" approach to measurement is ideal (Johnston, Brignall, & Fitzgerald, 2002). Provide narrative alongside indicators to put measures in context, always linking back to the desired outcomes and not focusing myopically on the indicators themselves (Ebrahim, 2005).

This section has added complexity to the simple impact-measurement story. It highlights the limitations of measurement in all but the simplest of situations, and it has presented unavoidable tensions and dysfunctions that are present in all performance-measurement systems. Best practice suggests assiduous use of indicators combined with human discretion in their interpretation.

Resources to help nonprofits select indicators:

Choosing and using indicators

S.M.A.R.T. and SICED Indicators

The Interconnectedness of Impact Measurement

This third section adds another level of complexity to performance assessment by noting that rarely is an organization free to determine its own measures and how to use them. Each organization's performance assessment is influenced and constrained by the actions of other organizations. *To fully understand impact measurement, we must understand it in the context of interconnected organizations, each using the measures for multiple purposes.*

Interconnectedness complicates the straightforward approach to performance assessment. Only in an imaginary world is impact measurement about a single organization (Ebrahim, 2005). In reality, nonprofits must negotiate their performance measurement with an interconnected web of funders, collaborators, and partners (O'Dwyer & Boosma, 2015; Thomson, 2011; Benjamin, 2008). Sometimes a more powerful organization, such as a funder, imposes measures. Sometimes a group of nonprofits chooses to align measures so that together they can be more powerful, such as a group of food banks that collaborate on advocacy work. At other times it has more to do with mundane management tasks, such as when a network of independently managed YMCAs collaborate to implement a provincewide project. Large organizations are increasingly conscious of how their measurement choices affect others. All this to say, there are many influences and constraints on a nonprofit's selection of objectives and measures, and these often extend well beyond its immediate stakeholders.

The influence of other organizations is not a one-time event. Best practice recommends that objectives and measures be revisited frequently and in dialogue with stakeholders. When a large funder revisits its causal model, there are implications for those it funds and those who seek funding. Indicators chosen by large coalitions can become *de facto* standards such that even organizations outside of the coalition find that they need to start tracking and reporting on those metrics. Impact measurement in less powerful nonprofits can be constantly at the whim of numerous uncoordinated, well-intentioned changes by the organizations around them.

The risk of promoting the straightforward impact-measurement approach is that, when we suggest that each organization set its own objectives, what we create is an ecosystem where only the powerful get to do so. The less powerful organizations spend their time supplying measurements to feed the bespoke approaches of their more powerful funders, partners, and collaborators and have no time to nurture their own measures (Coule, 2015). Nonprofits intending to follow the common practices (in dialogue with stakeholders) instead find themselves ensnared in hierarchical top-down systems to accommodate funder needs (Schmitz, Raggo, & Bruno-van Vijfejken, 2012).

Interconnectedness also complicates the problems that arise from the multiple purposes that measures must fulfil. This is because measures change roles as they move from one organization to the next. Governments, foundations, and membership bodies have their own accountabilities to the public (Newcomer, 2011; Ostrower, 2006). They too need to learn, advocate, and communicate. To do this they rely on the *accountability* measures provided by grantees. One organization's accountability is another organization's learning. Depending on what funders need to know, they may ask their grantees for different types of reporting, such as cost-benefit analysis, formative evaluation, or outcomes measurement (Carman, 2010). The different roles that measures must fulfil create additional tensions in an interconnected ecosystem.

Moreover, agreeing on metrics is challenging because different organizations in the ecosystem need different types of measures for different types of accountability and learning. Funders are accountable for community-level outcomes. A single implementing nonprofit is unlikely to have an attributable effect on those outcomes (Campbell, 2002), never mind have the capacity to measure that effect with contribution. This is a problem that "collective impact," discussed in the next section, has sought to address.

From "Anything Goes" to "One Right Way"

In response to the above coordination problems, many have argued that what we need is a *uniform* approach. Uniformity refers to a single set of processes and indicators to assess performance. If the problem is too much diversity in how organizations measure (Carman, 2010), uniformity might appear to be a reasonable solution. Uniformity is attractive because it appears to offer simplicity, comparability, and the ability to aggregate data.

Uniform approaches fall into four broad categories: indicator banks, collective impact, certifications, and commensuration.

Indicator banks are lists of indicators, organized by field or topic, that have been selected and agreed upon by a community of users (organizations, funders, etc). The indicators are then published so other organizations outside the initial group can use them. The objective of these standards is to get similar organizations measuring the same thing in the same way.

Indicator banks can be top-down, such as when funders require fundees to report on a list of prescribed indicators. This exacerbates the learning-accountability tension noted above. Uniform approaches can also undermine the organization's autonomy (Baur & Schmitz, 2012; Coule, 2015).

Alternatively, *indicator taxonomies* offer a flexible take-it-or-leave-it approach. Users select indicators from a "pool," only as are appropriate, while retaining the freedom to design their own measures. This is a preferable approach. Here are a few examples of this take-it-or-leave-it variety:

- Urban Institute's Outcome Indicators Project (2004) organizes outcome indicators into 14 fields of nonprofit work (for example adult education, performing arts, and prisoner re-entry).
- IRIS (2008) offers a catalogue of indicators for businesses, social enterprises, and nonprofits interested in assessing impact, particularly in the context of impact investing (see Chapter 14 by Harji and Hebb). The indicators cover 12 thematic areas, including agriculture, education, health, and housing. Indicators are numbered so they are easily identified. For example, "total permanent employees" is indicator O18869 and is defined as the "number of people employed by the organization as of the end of the reporting period." Many IRIS indicators are relevant to nonprofits.
- United Nations Sustainable Development Goals (UN SDGs) (2015) are an international agreement on 17 goals, each of which has several targets for which there are specific

indicators. Alliance 2030 is a network of Canadian organizations collaborating for the achievement and measurement of the UN SDGs.

- Global Reporting Initiative (GRI) (1997) is a set of indicators, but more than that, it is a standard for determining which economic, social, and environmental impacts an organization – particularly a for-profit corporation – should be measuring and reporting.

Collective impact eschews the notion that impact measurement should be determined by each organization for its own learning, management, accountability, and communication. Instead, it conceives of collectives of organizations coming together to tackle social problems with a shared vision, shared *uniform* measurement, and coordinating body, referred to as the "backbone organization" (Kania & Kramer, 2011; Cabaj & Weaver, 2016). Together these organizations learn, manage, account, and communicate. Together they set objectives and select indicators.

For example, the First 2000 Days Network is a collective-impact approach to early childhood development in six western provinces and northern territories in Canada. Partners aligned their measurement framework to allow "different organizations, with different approaches or strategies to be able to speak the same language and more easily work towards the same goal."

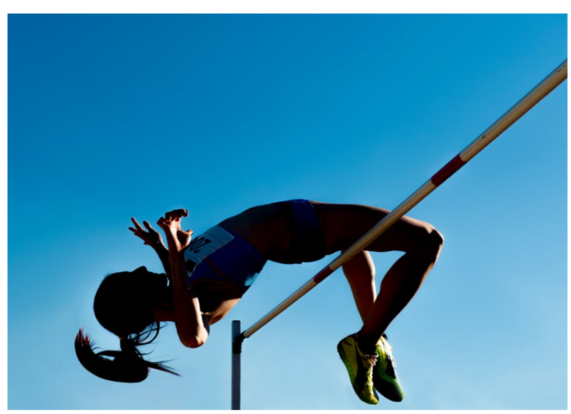
In essence, collective impact shifts impact-measurement basics (section one) to the collective rather than the organization. Although this has advantages, it risks impeding the learning and management systems within individual organizations (Baur & Schmitz, 2012; Coule, 2015). Further, it only partially solves part of the problem of interconnected organizations by shifting the problem of impact-measurement basics from organization-level problems to network-level problems.

Certifications and ratings are a different approach to uniformity (Carman & Fredericks, 2013; Feng, Neely, & Slatten, 2016). Both certification and rating have a series of performance criteria that organizations are evaluated against. Ratings are more like the long jump: every additional centimetre of impact counts. Certifications are more like the high jump: organizations clear the bar or they don't. With certifications, there is no difference between an organization that misses by a little and one that misses by a lot.

Figure 4: Ratings and certifications

Ratings are like the long jump: get as far as you can.

Certifications are like the high jump: clear the bar.



Examples of certifications are Canada Organic, FAIRTRADE, Certified B Corporation, and Imagine Canada's Standards Program (see Chapter 7, Phillips, Dougherty, & Barr). Charity Intelligence's star-rating of Canadian charities is akin to the long jump – a third-party rating system rather than certification, analogous to Charity Navigator and Candid (the merger of GuideStar and the Foundation Center) in the US (Sloan, 2008). It includes an assessment of the disclosure of impact information, but it does not evaluate impact due to limited public reporting.

B Lab's B Assessment and B Corp Certification is a combination of rating and certifying: it combines high jump with long jump. Organizations take the B Assessment, which generates a score (long jump). For certification, the bar is set at 80. Organizations can use their score or certification to communicate impact. Currently, there are more than 230 B Corps in Canada.

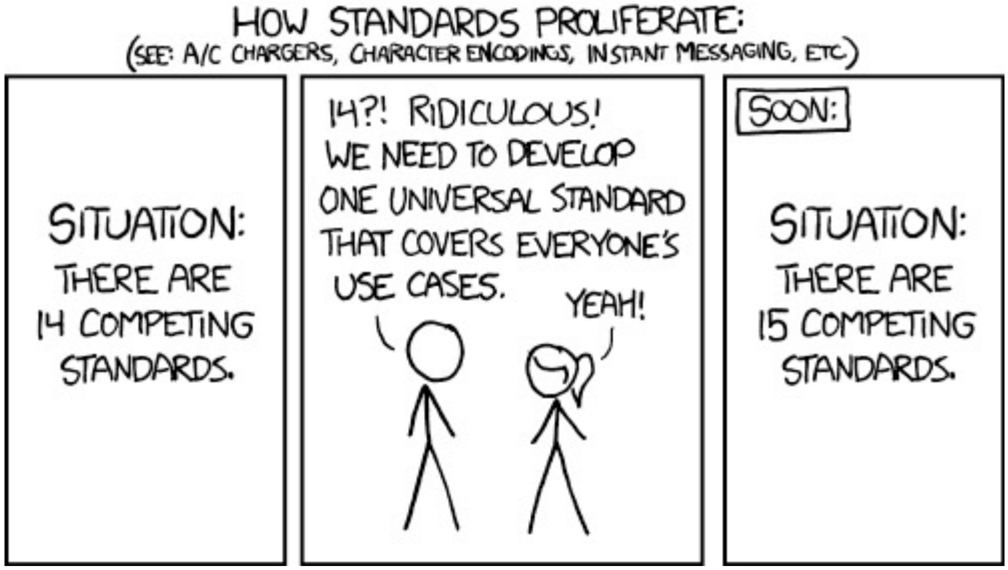
Commensuration converts all outcome indicators into a common metric, usually monetary. Once commensurate, it is easier to compare impacts. Cost-benefit and SROI are examples of commensuration. The SROI Guidebook (SROI, 2012: 11) advises against comparing the SROIs of organizations because variability in the process can cause similar outcomes to have dissimilar economic values. So, although it is *possible* to compare the SROI of a homeless shelter to that of a public library, in that one number will be higher than the other the *validity* of the comparison is questionable. Despite these words of caution, SROI and other commensuration techniques are often used to compare projects and organizations.

Limits and Challenges to Uniformity

There are several limits to uniformity (Ruff & Olsen, 2016). The individual organization's loss of autonomy – resulting in less ability to account, learn, manage, and communicate – has already been mentioned. In addition are the proliferation of standards and the danger of a single story. There have been many attempts to standardize charity and nonprofit performance measurement, dating as far back as the early 1900s. None of the early attempts are around today. The (proposed) standards of today are mostly recent endeavours.

A significant challenge with uniform standards is that if they are too lax or too rigorous, competing standards enter the market. This results in a proliferation of standards and a wide variation in quality and rigour that outsiders have a difficult time navigating (Reinecke, Manning, & von Hagen, 2012). It can be difficult for standards to remain meaningful, and the proliferation of standards becomes fodder for even further proliferation (see Figure 5).

Figure 5: How Standards Proliferate



(Source: https://xkcd.com/927/ Creative Commons licence)

Perhaps the greatest limitation of uniformity is the danger of creating a single story about *what matters* (Coule, 2015). All measurement systems value something and undervalue or ignore other things. Those choices are necessary for the system to be useful. However, those choices are dangerous when they are the be-all and end-all of what will count in society.

In sum, any uniform approach that deprives organizations of the ability to make measurement work for learning, improvement, and management is unlikely to have long-lasting traction (Timmermans & Epstein, 2010). If we are to have an impact-measurement standard, it must be a flexible standard (Timmermans & Epstein, 2010), meaning it must be effective for a single organization and a single purpose *and* provide a way of communicating, aggregating, and comparing (Ruff & Olsen, 2016). It must allow competing ways of thinking about what makes a "good" or effective organization (Brunsson, Rasche, & Seidl, 2012). It must also attend to the particularities of the situation rather than conformity to prescribed ends (Campbell, 2002; Coule, 2015). The Common Approach to Impact Measurement is one attempt at creating a flexible standard.

Nonprofits already know this. Although many are constrained by funder demands, there is evidence that nonprofits are finding space for flexibility and autonomy within those constraints by partially resisting the measurement obligations placed on them (Arvidson & Lyon, 2014). At present, that is the closest thing the sector has to a best practice for managing impact measurement in an interconnected ecosystem.

Conclusion: Takeaways for Canadian Organizations

This section endeavours to translate the above content into practical steps that Canadian nonprofit leaders, and their public- and private-sector partners, can take today to improve impact measurement in their organizations.

Seven Tips for More Effective Impact Measurement and Use

1. ***Go for "good enough" when measuring impact.*** Doing a good job measuring outputs and near-term outcomes can be enough when combined with other low-cost steps. Use evidence-based assumptions to extrapolate outputs to impact. Compare changes across time (before and after). Organize program work in mini-trials or experiments and compare across sites to learn what works best. This can be done with a few key performance indicators. (See the Robin Hood Foundation for an example.)
2. ***Engage stakeholders.*** Not only will your organization be better informed, but a measurement approach that is well connected to what community members want and need is a powerful anchor. It can be useful in persuading others to accept your measures rather than adapting to theirs. A cautionary note: some stakeholder engagement can be invasive and create burdens on respondents. Respect the privacy and time of program participants.
3. ***Use common indicators as much as possible.*** Using a causal model and common practices should suggest relevant outputs and outcomes. Check to see if some of the indicator taxonomies (e.g. IRIS, SDGs, Urban Institute, GRI) have indicators to address those output and outcomes, and use them if they do. Again, these are powerful anchors that will help to protect your organization's performance assessment from external influence. Do not change your organization's objectives to match an indicator!
4. ***Use technology to harness the data the organization is already tracking.*** In my experience, most charities and social enterprises track sufficient data but in a disorganized way: they are either on paper or in files on individual computers. Disparate data, whether quantitative or qualitative, are difficult to use when writing reports and grant applications.
5. ***Gather a range of indicators to mitigate the tensions and dysfunctions of measurement, but don't gather more data than you are going to use.*** And, especially if you are a funder, don't ask grantees for impact measures that don't figure into your decision-making and accountability systems. More performance measurement is not always a good thing (Ebrahim, 2005).
6. ***The benefits of transparency and accountability have their limits.*** It is okay to keep some metrics private. Accountability is seen as both a mechanism and a virtue (Messner, 2009; Roberts, 2009). Nonprofits may feel that they must share everything to be more virtuous. It is okay to abide by the mechanism and keep some measures private for real learning out of the spotlight.
7. ***Enjoy it!*** Measurement is kind of fun. Most nonprofit and social enterprise leaders are passionate about the work they do, and measurement is an ideal weekly task to let you really think big about what you are doing, why, and how you can do better.

References

Alexander, J., Brudney, J.L., & Yang, K. (2010) Introduction to the symposium: Accountability and performance measurement: The evolving role of nonprofits in the hollow state. *Nonprofit and Voluntary Sector Quarterly*, 39(4), 565–570.

Arvidson, M. & Lyon, F. (2014) Social impact measurement and non-profit organisations: Compliance, resistance, and promotion. *Voluntas*, 25, 869–886.

Baker, M. & Schaltegger, S. (2015) Pragmatism and new directions in social and environmental accountability research. *Accounting, Auditing and Accountability Journal*, 28(2), 263–294.

Baur, D. & Schmitz, H.P. (2011) Corporations and NGOs: When accountability leads to co-optation. *Journal of Business Ethics*, 106(1), 9–21.

Benjamin, L.M. (2008) Account space: How accountability requirements shape nonprofit practice. *Nonprofit and Voluntary Sector Quarterly*, 37(2), 201–223.

Brignall, S. & Modell, S. (2000) An institutional perspective on performance measurement and management in the "new public sector." *Management Accounting Research*, 11(3), 281–306.

Brunsson, N., Rasche, A., & Seidl, D. (2012) The dynamics of standardization: Three perspectives on standards in organization studies. *Organization Studies*, 33(5-6), 613–632.

Burchell, S., Clubb, C., Hopwood, A., Hughes, J., & Nahapiet, J. (1980) The roles of accounting in organizations and society. *Accounting, Organizations and Society*, 5(1), 5–27.

Cabaj, M. & Weaver, L. (2016) *Collective impact 3.0: An evolving framework for community change*. Waterloo, ON: Tamarack Institute.

Campbell, D. (2003) Outcomes assessment and the paradox of nonprofit accountability. *Nonprofit Management and Leadership*, 12(3), 243–259.

Campbell, D.T. (1979) Assessing the impact of planned social change. *Evaluation and Program Planning*, 2(1), 67–90.

Carman, J.G. (2010) The accountability movement: What's wrong with this theory of change? *Nonprofit and Voluntary Sector Quarterly*, 39(2), 256–274.

Carman, J. & Fredericks, K. (2013) Nonprofits and accreditation: Exploring the implications for accountability. *International Review of Public Administration*, 18(3), 51–68.

Chen, K.K. (2013) Storytelling: An informal mechanism of accountability for voluntary organizations. *Nonprofit and Voluntary Sector Quarterly*, 42(5), 902–922.

Chenhall, R.H., Hall, M., & Smith, D. (2013) Performance measurement, modes of evaluation and the development of compromising accounts. *Accounting, Organizations and Society,* 38(4), 268–287.

Chenhall, R.H., Hall, M., & Smith, D. (2017) The expressive role of performance measurement systems: A field study of a mental health development project. *Accounting, Organizations and Society*, 63, 60–75.

Connolly, C. & Hyndman, N. (2013) Towards charity accountability: Narrowing the gap between provision and needs? *Public Management Review,* 15(7), 945–968.

Coule, T.M. (2015) Nonprofit governance and accountability: Broadening the theoretical perspective. *Nonprofit and Voluntary Sector Quarterly*, 44(1), 75–97.

Dhanani, A. & Connolly, C. (2012) Discharging not-for-profit accountability: UK charities and public discourse. *Accounting, Auditing and Accountability Journal*, 25(7), 1140–1169.

DiMaggio, P. (2001) Measuring the impact of the nonprofit sector on society is probably impossible but possibly useful. In P. Flynn & V.A. Hodgkinson (Eds.), *Measuring the impact of the nonprofit sector*, 249–272. Springer.

Earl, S., Carden, F., & Smutylo, T. (2001) *Outcome mapping: Building learning and reflection into development programs*. Ottawa: International Development Research Centre.

Ebrahim, A. (2005) Accountability myopia: Losing sight of organizational learning. *Nonprofit and Voluntary Sector Quarterly*, 34(1), 56–87.

Ebrahim, A. & Rangan, K.V. (2010) Putting the brakes on impact: A contingency framework for measuring social performance. *Academy of Management Proceedings*, (1).

Feng, N.C., Neely, D.G., & Slatten, L.A.D. (2015) Accountability standards for nonprofit organizations: Do organizations benefit from certification programs? *International Journal of Public Administration*, 39(6), 470–479.

Global Reporting Initiative. (2016) Consolidated set of GRI sustainability reporting standards 2016. Retrieved from https://www.globalreporting.org/standards/

Hall, M. (2008) The effect of comprehensive performance measurement systems on role clarity, psychological empowerment and managerial performance. *Accounting, Organizations and Society*, 33(2-3), 141–163.

Hansen, H., Klejnstrup, N.R., & Andersen, O.W. (2013) A comparison of model-based and design-based impact evaluations of interventions in developing countries. *American Journal of Evaluation*, 34(3), 320–338.

Herman, R.D. & Renz, D.O. (1999) Theses on nonprofit organizational effectiveness. *Nonprofit and Voluntary Sector Quarterly*, 28(2), 107–126.

Johnston, R., Brignall, S., & Fitzgerald, L. (2002) "Good enough" performance measurement: A trade-off between activity and action. *Journal of the Operational Research Society*, 53(3), 256–262.

Kania, J. & Kramer, M. (2011) Collective impact. *Stanford Social Innovation Review*. Retrieved from http://ssir.org/articles/entry/collective_impact

Kaplan, R.S. (2001) Strategic performance measurement and management in nonprofit organizations. *Nonprofit Management and Leadership*, 11(3), 353–370.

Keevers, L., Treleaven, L., & Sykes, C. (2012) Made to measure: Taming practices with results-based accountability. *Organization Studies*, 33, 97–120.

Liket, K.C. & Maas, K. (2013) Nonprofit organizational effectiveness: Analysis of best practices. *Nonprofit and Voluntary Sector Quarterly*, 44(2).

McLaughlin, J.A. & Jordan, G.B. (1999) Logic models: A tool for telling your program's performance story. *Evaluation and Program Planning*, 22, 65–72.

Mayne, J. (2001) Addressing attribution through contribution analysis: Using performance measures sensibly. *The Canadian Journal of Program Evaluation*, 16(1), 1–24.

Melnyk, S., Stewart, D., & Swink, M. (2004) Metrics and performance measurement in operations management: Dealing with the metrics maze. *Journal of Operations Management*, 22(3), 209–218.

Messner, M. (2009) The limits of accountability. *Accounting, Organizations and Society*, 34(8), 918–938.

Mitchell, R.K., Wood, D.J., & Agle, B. (1997) Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts. *Academy of Management Review,* 22(4), 853–886.

Mudaliar, A., Pineiro, A., Bass, R., & Dithrich, H. (2017) The state of impact measurement and management practice. Retrieved from https://thegiin.org/research/publication/imm-surveyh

Newcomer, K.E. (2016) Tracking and probing program performance: Fruitful path or blind alley for evaluation professionals? *American Journal of Evaluation*, 22(3), 337–341.

Oakes, L. & Young, J. (2008) Accountability re-examined: Evidence from Hull House. *Accounting, Auditing and Accountability Journal*, 21(6), 765–790.

O'Dwyer, B. & Boomsma, R. (2015) The co-construction of NGO accountability. *Accounting, Auditing and Accountability Journal*, 28(1), 36–68.

O'Dwyer, B. & Unerman, J. (2008) The paradox of greater NGO accountability: A case study of Amnesty Ireland. *Accounting, Organization and Society*, 33(7-8), 801–824.

Phillips, S.D. & Carlan, V. (2018) On impact: Emerging challenges of evaluation for Canada's nonprofit sector. In K.E. Seel (Ed.), *The management of nonprofit and charitable organizations in Canada* (3rd ed.). Toronto: LexisNexis.

Poole, D.L., Nelson, J., Carnahan, S., Chepenik, N.G., & Tubiak, C. (2000) Systems in nonprofit agencies: The program accountability quality scale (PAQS). *American Journal of Evaluation*, 21(1), 15–26.

Reinecke, J., Manning, S., & von Hagen, O. (2012) The emergence of a standards market: Multiplicity of sustainability standards in the global coffee industry. Organization Studies, 33(5–6), 791–814.

Roberts, J. (2009) No one is perfect: The limits of transparency and an ethic for "intelligent" accountability. *Accounting, Organizations and Society,* 34, 957–970.

Ruff, K. & Olsen, S. (2016) The next frontier in impact measurement isn't measurement at all: Why we need skilled analysts to improve social capital markets. *Stanford Social Innovation Review*. Retrieved from https://ssir.org/articles/entry/the_next_frontier_in_social_impact_measurement_isnt_measurement_at_all

Saxton, G.D. & Guo, C. (2011) Accountability online: Understanding the web-based accountability practices of nonprofit organizations. *Nonprofit and Voluntary Sector Quarterly,* 40(2), 270–295.

Schmitz, H.P., Raggo, P., & Bruno-van Vijfeijken, T. (2012) Accountability of transnational NGOs: Aspirations vs. practice. *Nonprofit and Voluntary Sector Quarterly*, 41(6), 1175–1194.

Sloan, M.F. (2008) The effects of nonprofit accountability ratings on donor behavior. *Nonprofit and Voluntary Sector Quarterly*, 38(2), 220–236.

SROI Network. (2012) A guide to social return on investment. Retrieved from http://www.socialvalueuk.org/resource/a-guide-to-social-return-on-investment-2012h

Thomson, D.E. (2010) Exploring the role of funders' performance reporting mandates in nonprofit performance measurement. *Nonprofit and Voluntary Sector Quarterly*, 39(4), 611–629.

Thomson, D.E. (2011) The role of funders in driving nonprofit performance measurement and use in strategic management. *Public Performance & Management Review*, 35(1), 54–78.

Timmermans, S. & Epstein, S. (2010) A world of standards but not a standard world: Toward a sociology of standards and standardization. *Annual Review of Sociology,* 36(1), 69–89.

Wardle, L. (2017) The universal standards for social performance management implementation guide. Social Performance Task Force.

Weiss, C.H. (1997) Theory-based evaluation: Past, present, and future. *New Directions for Evaluation*, 1997(76), 41–55.

White, H. (2010) A contribution to current debates in impact evaluation. *Evaluation*, 16(2), 153–164.

Wimbush, E., Montague, S., & Mulherin, T. (2012) Applications of contribution analysis to outcome planning and impact evaluation. *Evaluation*, 18(3), 310–329.

W.K. Kellogg Foundation. (2004) *Foundation logic model development guide*. Retrieved from https://www.wkkf.org/resource-directory/resources/2004/01/logic-model-development-guide

Biography

Kate Ruff, Carleton University

Kate Ruff is an assistant professor at the Sprott School of Business and the co-director of the Carleton Centre for Community Innovation. Kate researches how organizations measure and report their social results. She is currently leading a project called the Common Approach to Impact Measurement.