



Measuring Social Impact in Arts and Cultural Organizations: A Review

Prepared for
The Dundas Valley School of Art

In

September 2020

By

Umair Majid
Jasmine Dzerounian
Sureka Pavalagantharajah
Haley Hutchison
Kelly Wu

Contents

- Executive Summary 3
- Introduction..... 5
- Background..... 7
 - About the Dundas Valley School of Art..... 7
 - Performance Measurement in Non-Profit Organizations 7
 - Measuring Social Impact in Arts and Cultural Organizations 8
- Methodology 9
 - Academic Database Searching..... 9
 - Grey Literature Searching..... 9
 - Data Extraction and Synthesis 10
 - Limitations 10
- Findings 11
 - Determining what to measure 11
 - Logic Models 12
 - Social Impact Measurement Methods..... 13
 - Experimental and Implicit Designs..... 13
 - Cost-Benefit Analysis and Social Return on Investment..... 14
 - Stakeholder Involvement 16
 - Indicators and Indicator Selection 16
 - Barriers to Measuring Social Impact 18
 - Facilitators for Measuring Social Impact..... 19
- Recommendations 20
 - Recommendation 1: Determine the purpose of social impact measurement..... 20
 - Recommendation 2: Develop a robust social impact measurement framework informed by lessons from diverse disciplines and context, but that is unique to the needs of the DVSA 21
 - Recommendation 3: Consider including health indicators to measure social impact 21
 - Recommendation 4: Involve stakeholders (users, funders, and art professionals) in the development and implementation of social impact measurement 21
 - Recommendation 5: Strengthen organizational capacity for high-quality social impact measurement 22
- Conclusion 22
- Bibliography 23
- Appendices..... 28

Appendix A: Search Strategy.....	28
Appendix B: Descriptive Characteristics of Academic Literature	29
Appendix C: Grey Literature Resources.....	58

Executive Summary

The primary objective of this research was to review the academic and grey literature to identify lessons, considerations, methods, and approaches for measuring social impact in arts and cultural organizations, specifically for the Dundas Valley School of Art (DVSA). The research question for this report was: How might the DVSA set about to measure and report on their social impact?

We conducted an academic database and grey literature searches to identify frameworks, indicators, and other tools for measuring social impact in non-profit organizations including arts and cultural organizations. Our report is divided into the following sections: (1) determining what to measure, (2) social impact measurement methods, (3) stakeholder involvement, (4) indicators and indicator selection, (5) barriers to measuring social impact, (6) facilitators to measuring social impact.

In the first section (determining what to measure), we discuss common frameworks that have been used to measure the social impact of non-profit organizations, including Cost Benefit Analysis, Social Return on Investment, and Logic Models. We describe the differences between outputs (i.e., countable goods and services produced), outcomes (i.e., changes to intended beneficiaries as a result of organizational activities), and impact (i.e., contributions of an organization's activities to its overall social mission). We also use three examples from existing arts and cultural organizations to describe how they conceptualized and measured their social impact.

In the second section (social impact measurement methods), we identify basic approaches to measuring social impact. We differentiate between experimental and implicit or quasi-experimental study designs. The main difference between these study designs lies in the nature of the control group. We discuss the advantages and disadvantages of both approaches that organizations might consider before deciding which one is more appropriate for their goals.

In the third section (stakeholder involvement), we discuss the importance of stakeholder involvement in the design and administration of social impact measurement. We highlight potential approaches to involving diverse stakeholders using examples from a Canadian arts education organization.

In the fourth section (indicators and indicator selection), we identify three categories of indicators relevant for social impact measurement: output-focused, outcome-focused, and mission-focused. We also discuss several considerations for selecting indicators, including the characteristics of strong and weak indicators. Finally, we describe two consensus-based approaches (Delphi survey and nominal group technique) that organizations can use to prioritize the most relevant indicators for measuring their social impact.

In the fifth section (barriers to measuring social impact), we explain the three most common barriers to measuring social impact: organizational culture, resource

constraints, and accountability and relationships. Similarly, in the final section (facilitators to measuring social impact), we describe the three most important facilitators for measuring social impact: community outreach, data, and funding.

Based on our findings, we propose five recommendations for the DVSA as steps towards measuring their social impact:

- (1) determine the purpose of social impact measurement
- (2) develop a robust social impact measurement framework informed by lessons from diverse disciplines but that is unique to the needs of the DVSA
- (3) consider including health indicators to measure social impact
- (4) involve stakeholders (users, funders, and art professionals) in the development and implementation of performance measurement
- (5) strengthen organizational capacity for high-quality performance measurement

Introduction

Overview: This report was prepared by Research Associates from the McMaster Research Shop at the request of the Dundas Valley School of Art (DVSA). This research includes a review of academic and grey literature to identify methods for measuring social impact performance in non-profit organizations, particularly those that provide community-based arts and cultural services.

Context: Guided by frameworks like the Canadian Index of Well-being, civil sector funding agencies are increasingly looking for non-profit organizations to clearly and transparently communicate their social impact. Individual organizations are responsible for measuring and reporting on their own social impact; however, social impact is difficult to define and quantify (Hager, 2001; Lall, 2019).

More than just an art school, the DVSA, in addition to their regular programming, provides art-based workshops in partnership with over 20 social service agencies and programs in the region. Through its community outreach and therapeutic art programming, the DVSA believes it is helping to support and strengthen their community. However, developing a process to collect performance metrics for DVSA's programs will help communicate their social value and make the case that an art school is an important community service. If the DVSA can better communicate its impact, they will be better positioned to secure funding to continue their work.

Scope: The guiding question for this project is, "how can the DVSA measure and report on their social impact?". This report draws on academic and grey literature on measuring performance and social impact in non-profit organizations, particularly community-based arts and cultural services.

Terms: In this report, we will use the term *performance measurement* to indicate a process for assessing how well an organization is achieving its goals, which may include financial, administrative, and social indicators. Where our discussion is specific to *social impact*, we will use this term.

Overview: The report is organized in the following way:

- **Background:** This section introduces performance measurement in non-profit organizations, the challenges with measuring social impact, the rationale for this research, and a description of the Dundas Valley School of Art
- **Methodology and Limitations:** This section describes the purpose and approach of our database and grey literature approaches. We also discuss the limitations with our research.
- **Findings:** This section synthesizes our findings, which we have divided into the following sub-categories:
 - **Determining What to Measure:** We describe the characteristics of frameworks that can be used to conceptualize social impact. We identify logic models as an appropriate framework to use for measuring social impact in the context of arts and cultural organizations.

- Social Impact Measurement Methods: We identify basic approaches to measure social impact: experimental and implicit designs. We also discuss Cost Benefit Analysis and Social Return on Investment as common approaches to measuring social impact.
- Stakeholder Involvement: We discuss the importance of stakeholder involvement in the design and implementation of social impact measurement, as well as ways to involve them throughout the impact measurement process.
- Indicators and Indicator Selection: We identify categories of indicators relevant for social impact measurement: output-focused, outcome-focused, and mission-focused. We describe the characteristics of high-quality indicators and discuss two techniques for prioritizing the most relevant indicators.
- Barriers to Measuring Social Impact: We explain the three most common barriers to measuring performance and social impact: organizational culture, resource constraints, and accountability and relationships.
- Facilitators to Measuring Social Impact: We describe the three most important facilitators of measuring performance and social impact: community outreach, data, and funding.
- **Recommendations**: Based on our findings, we recommend five steps for the Dundas Valley School of Art: (1) determine the purpose of social impact measurement; (2) develop a robust social impact measurement framework informed by lessons from diverse disciplines but that is unique to the needs of the DVSA; (3) consider including health indicators to measure social impact (4) involve stakeholders (users, funders, and art professionals) in the development and implementation of performance measurement; and (5) strengthen organizational capacity for high-quality performance measurement.

Background

About the Dundas Valley School of Art

The Dundas Valley School of Art (DVSA), established in 1964, is a community arts institution located in Dundas, Ontario. While primarily serving Dundas, the DVSA also serves other areas of Hamilton. The DVSA offers professional art services and classes such as painting, photography, pottery, sculpting, and printmaking to community members from all walks of life. Shortly after opening, the DVSA expanded to offer a network for local artists to sell their work through auctions and showcases. Though the DVSA is a local organization, it is also an active member of a community of art schools across Canada.

Beyond teaching art to children and adults, the DVSA also provides art-based programming for health and well-being. The DVSA works closely with over twenty social service agencies, including the Down Syndrome Association, Big Brothers and Big Sisters, and the Autism Society, to provide programs and art therapy workshops to those who may benefit from artistic expression. Additionally, long-term care facilities and retirement residences in the city have worked closely with the DVSA to facilitate therapeutic art workshops. The DVSA has also partnered with organizations that serve marginalized populations like previously incarcerated individuals and female-identifying individuals.

Performance Measurement in Non-Profit Organizations

In the last two decades, funding agencies have increasingly required organizations to report on their performance. Performance measurement is the ongoing process of collecting, analyzing, and reporting on information relating to an organization's activities and impact (Arena, 2015). Measuring performance is especially important for non-profit organizations (NPOs) because of their reliance on external funding. NPOs must demonstrate that they are using their resources, including funding from community foundations and government agencies, to accomplish what they set out to do. In addition, measuring performance allows NPOs to reflect and re-evaluate programs in order to better serve their communities.

Originally, performance measurement was designed for for-profit organizations as a way to document and communicate the efficiency or effectiveness of their activities (Arena, 2015). Some performance measurement systems also consider the long-term success of a business based on the indicators of service quality and customer satisfaction. These performance measurement indicators are usually collected through semi-structured interviews and surveys with staff and service users (Arena, 2015). Performance measurement involves developing and selecting appropriate financial and non-financial indicators, ongoing data collection and analysis on data related to these

indicators and reporting on performance indicators to improve the quality of services or acquire additional resources to support existing services (Harbour, 2017).

Measuring Social Impact in Arts and Cultural Organizations

Measuring social impact has become increasingly important for NPOs and their donors and funders. Measuring impact often includes the organization's influence or effect on its social, environmental, or economic surroundings. However, social impact focuses on the organization's effect on a pressing social challenge (Ibrisevic, 2018). The Social Impact of the Arts (1993) describe social impact of the arts as "those effects that go beyond the artefacts and the enactment of the event or performance itself and have a continuing influence upon, and directly touch, people's lives" (Landry et al., 1993, p. 50). In the case of arts and cultural organizations, social impact includes the effect of arts-based activities on social issues including equality, livelihoods, mental health, poverty, justice, social inclusion, and education.

The current landscape for funding in Canada emphasizes investing in initiatives that have a social impact. For example, the Ontario Trillium Foundation (OTF) prioritizes organizations committed to improving the well-being of their communities. They cite the Canadian Index of Well-Being that measures the quality of life of Canadian communities (Ontario Trillium Foundation, 2020). Arts and culture organizations are often funded on the basis of their social value. For instance, the OTF asserts that arts organizations have the capacity to inspire and connect communities (Ontario Trillium Foundation, 2020). This emphasis on social change through the arts demands greater emphasis on accountability and transparency in terms of social impact and performance for arts and culture NPOs (Polonsky and Grau, 2011).

The pressure on the arts and cultural organizations to justify its value is compounded with increased NPO competition for scarce funds. Accurate measurement of social impact for arts and cultural organizations and their interventions is a source of competitive advantage. Challenges that arts and cultural organizations face in assessing social impact is that there are no standard performance measurements and the lack of robust evaluation tools and systematic evidence of the impact of arts projects or cultural services (Polonsky and Grau, 2011). Significant limitations in current methodologies to measure social impact include small sample surveys, reliance on self-reported measures, use of single sample case studies, and a lack of data on processes (Jermyn, 2001). There are opportunities for arts and cultural organizations to address the barriers to measuring social impact including measuring individual confidence and self-esteem outcomes, educational impacts, local economic impact and regeneration, health promotion, increased social cohesion, community development, community empowerment, social inclusion, a sense of local identity, and sustainable development (Coalter, 2001).

Methodology

For this research, we identified academic research and grey literature describing performance measurement frameworks or systems that could help DVSA measure and report on their social impact, including the experiences and perspectives of practitioners. We focused on performance measurement for non-profit organizations and excluded research on the for-profit sectors due to different organizational structures, values, and goals.

Academic Database Searching

We conducted a database search in the Arts & Humanities and Social Sciences Index of Web of Science using keywords and subject headings related to non-profit organizations, social enterprises, and performance measurement (see Appendix A for search strategy). We screened relevant citations via Covidence, prioritizing articles that discussed performance measurement frameworks or systems, examples of how a non-profit organization or social enterprise measured performance, the conceptualization and operationalization of indicators that organizations used to measure performance, the perspectives of stakeholders towards performance measurement, and any of these topics pertaining to measuring social impact. We also retrieved studies on performance measurement in arts and cultural organizations. We included relevant articles from all sectors including recreation, health care, and human services. However, we only included articles from Australia, Canada, New Zealand, the United Kingdom, and the United States, because these countries are similar in their funding environment for NPOs.

We screened the title and abstracts of 528 studies, finding 75 potentially relevant studies. We used this literature to inform our findings and recommendations sections. While not all studies were used directly in this report, we found five Canadian studies, three of which were conducted in the context of arts and cultural organizations (Chahine, 2012; Reid and Karrambayya, 2009; Turbide and Laurine, 2009). See Appendix B for a summary of descriptive information of all 75 articles from the academic literature.

Grey Literature Searching

We surveyed Google.com and Google.ca using keywords similar to our database search. Our goal was to use grey literature to complement the academic literature. We were most interested in learning from non-profit organizations or social enterprises (including arts and cultural organizations) who have well-developed performance measurement systems for social impact.

Four researchers conducted 13 rounds of grey literature searching using Google.ca and Google.com, resulting in 175 relevant hits. These webpages and documents were reviewed for relevance. We reviewed all hits and decided on 19 webpages and

documents that were most relevant for our research questions. We described these resources in Appendix C.

Data Extraction and Synthesis

We extracted descriptive information from included material, including the following: country of publication, type of article or study design, and topic. We conducted summary statistics on these characteristics to give an overview of the characteristics of included literature.

Using thematic analysis, five researchers reviewed the findings and discussion sections of five articles to develop a list of themes, ideas, and concepts pertaining to reporting and communicating social impact. We used these themes to develop templates that categorized relevant findings from our analysis. We extracted relevant findings from all 75 articles into the following templates: name, purpose, and dimensions of performance measurement frameworks; lessons for performance measurement; indicators or methods to measure performance; and communication and reporting practices. After we extracted relevant findings from all included studies, we reviewed the templates to develop narrative summaries that provided a summary of each theme. It is important to note that when reviewing the templates, we did not include data from all 75 studies, as some data was not relevant for the goals of this review. We have included the descriptive and study characteristics of all 75 studies in the appendix in case other articles are of interest to readers.

Limitations

Though we found a diversity of academic and grey literature, the information is limited in the following ways:

- We were only able to locate only five academic studies published in Canada (Bourgeois et al., 2015; Chahine, 2012; Mackinnon and Stephens, 2010; Reid and Karrambayya, 2009; Turbide and Laurine, 2009). Compared to the quantity of included literature, this is a major limitation that might affect how well our findings can be applied to NPOs in Canada. More research is needed on measuring performance and social impact in Canadian NPOs.
- Most academic studies did not provide demographic information of the communities and individuals of interest. Since Hamilton and nearby communities are diverse, we would be careful in applying our findings directly to the Hamilton context.
- Due to time constraints, we only searched one academic database for relevant articles. While the database we chose was most relevant to our research question, it is probable that we missed important articles that could offer additional insight on how to measure and report on social impact in arts and cultural organizations.
- While we found extensive discussion on performance measurement in non-profit organizations, most included studies focused on administrative, human resource,

capacity, and financial performance. We found minor discussion of social impact in the NPO literature, particularly for arts and cultural organizations. We also found little mention of the methods or tools that NPOs can use to measure social impact. This characteristic made it challenging for us to provide useful advice for measuring social impact.

Findings

Determining what to measure

Social impact measurement requires first selecting what you want to measure. When measuring social impact, or “public good”, Liket et al. (2014) distinguishes between *mission-related impact* and *public good impact*. Mission-related impact refers to the impact the NPO has generated based on their mission statement. Public good impact is defined as the total intended, unintended, positive, and negative changes to individuals, organizations, the environment, communities, and institutions. Evaluating public good effectiveness is necessary for organizations who want to have an impact beyond their organization’s mission or mandate.

Developing a social impact measurement system also requires assessing different *levels* of effectiveness. The literature distinguishes between outputs, outcomes, and impact (Liket et al., 2014; Lee and Nowell, 2015):

- *Outputs* are countable goods and services (or otherwise “things”) that are produced by an NPO, for instance the number of seniors who participate in therapeutic arts programming.
- *Outcomes* are changes to the intended beneficiaries as a result of a program’s activities, such as reduced feelings of social isolation for seniors participating in therapeutic arts programming.
- *Impact* is the extent to which an organization contributes to their mission or public good. The evaluation literature sometimes calls this the “long-term outcome”. For instance, the intended impact or long-term outcome of a student food bank might be increased food security for post-secondary students.

For NPOs, determining impact requires addressing what “dimensions” of public good they are contributing to. Lee and Nowell (2015) discuss *public value accomplishment*, which refers to community outcomes and broader benefits to society. While for-profit organizations use profit maximization to measure their success, public value accomplishment is more appropriate to use as an NPO, as it highlights the organization’s importance and its contribution to society.

Public value can include a number of measures, including quality of life, wellbeing, happiness, social inclusion, safety, equality, innovation, advocacy, individual expression, and citizen engagement in the community (Lee and Nowell, 2015). As an example, one could measure public value accomplishment in the form of health benefits

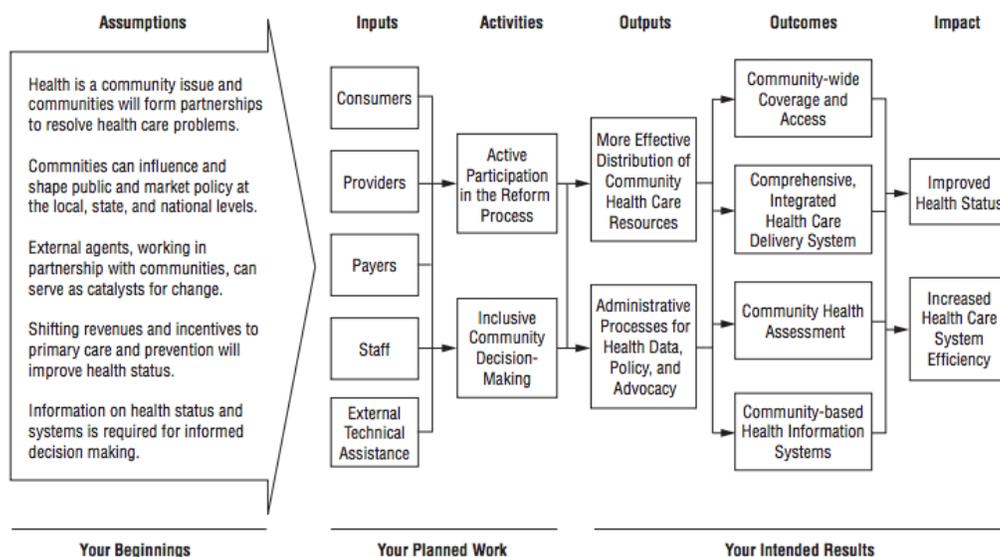
local patrons of a community NPO experience over a span of four months. *Social ambition* is related to public value accomplishment, because NPOs' ultimate goal is to achieve their social ambitions that they define in their organizational mission. Since many NPOs' mission includes a cause that accomplishes public value, when NPOs achieve their social ambition, they also accomplish public value. A number of public frameworks exist to help identify an NPOs impact, and hence what to measure. For instance, an organization could measure the extent to which it contributes to a UN sustainable development goals or one (or more) of the eight dimensions of the Canadian Index of Wellbeing.

Logic Models

A logic model (also called a “theory of change”) can be used to link an organization’s inputs, activities, and different effect levels (i.e., outputs, outcomes, and social impact). Funders appreciate logic models because they provide the ability to fund discrete program components. Sound logic models are achieved through comprehensive discussions amongst key stakeholders that result in a high degree of coherence between a NPO’s activities and intended impact (Lynch-Cerullo and Cooney, 2011).

The W.K. Kellogg Foundation is a national NPO in the United States that prioritizes services for vulnerable children in order to build more equitable communities. They provide a logic model development guide that outlines the process of how a NPO can develop their logic model. The guide outlines examples of logic models, such as the Comprehensive Community Health Models of Michigan. This model outlines assumptions made at the beginning of the process, as well as inputs, activities, outputs, outcomes, and impacts. The figure below is a copy of the logic model (W.K. Kellogg Foundation, 2004, p. 11).

Figure 1: Example of a theory of change from W.K. Kellogg Foundation



The literature on evaluation in arts-based organizations and art schools is fairly limited. With regard to published evaluations, the Yukon School of Visual Arts (SOVA) conducted an evaluation in 2014 to look at their long-term outcomes (Kishchuk, 2014). They used a logic model to evaluate their long-term impact in the following three areas:

1. Training individuals to be use a variety of visual arts practices influenced by Yukon's environment, culture, and historical context
2. Contributions to cultural industries such as First Nation visual art in Yukon and Canada
3. Establishing their role as a foundation-level visual arts educator in Yukon.

The logic model was informed by SOVA's March 2012 Strategic Plan and was reviewed and approved by SOVA's Governance Council in September 2014. Data that they used to conduct their evaluation included the following:

- Administrative data and documents
- Financial data for the program
- Enrolment and completion data
- Exit surveys
- Interviews with current and former staff
- Interviews with program graduates
- Interview with Governance Council members
- Interview with two representatives from the funding agency (Yukon Advanced Education) and another stakeholder (Kishchuk, 2014).

The Ontario Arts Council, while not an art school, provides funding to initiatives aimed at improving social impact (Ontario Arts Council, n.d.). Some of the social impact they have considered include the following: increased talent and skills among Ontarians, particularly youth; healthier places and populations; community vitality and quality of life; greater participation and inclusion for all in community life; and improved public safety and security. The indicators they have used in their evaluation plan include, but are not limited to:

- Percentage of Ontario districts where resident artists or arts organizations received Ontario Arts Council funding
- Number of artists and arts professionals who attend training and skill development events or activities provided by the Ontario Arts Council
- Number of public activities or events held by Ontario Arts Council-funded artists and organizations in their home communities (Ontario Arts Council, n.d.).

Social Impact Measurement Methods

Experimental and Implicit Designs

After an organization has determined what they want to measure, the next step is to determine *how* they are going to measure performance. The literature provides many different ways to evaluate outputs, outcomes, and impact. Some program evaluation

designs may only measure outcomes at the end of a program. However, comparing participant experiences, attitudes, or beliefs before and after participating in a program or service will enable an organization to more accurately measure the impact (Grembowski, 2016).

There are two primary types of evaluation designs: an experimental design or an implicit (also known as “quasi-experimental”) design (Grembowski, 2016). *Experimental designs* involve selecting a portion of individuals from a community randomly to participate in a program, while the rest do not participate. Those who are not selected to participate are referred to as the “control group” and those that were selected to participate in the program are the “experimental group.” The evaluator will compare the controls’ outcomes to those of the experimental group, in order to determine the effect of the program on participants. In some cases, it is not feasible to perform experimental designs because it requires data collection from program participants and a randomly selected portion of the target community. In such cases, an implicit design may be used instead. In implicit designs, the evaluation does not have a “control group.” This does not necessarily mean that there is no comparison group at all. While some implicit designs may only look at the outcomes of program participants, other implicit designs may compare program participants to another group that did not participate in the program and was also not randomly selected (Grembowski, 2016). An example of a comparison where the selection is not random would be comparing the outcomes of two similar arts-based initiatives with populations that might differ on key demographic characteristics such as age and sex. In this case, we cannot say that the two groups are randomly selected because there was no method of “randomization,” such as a coin flip or an online random number generator to assign them to each clinic.

The drawback with implicit designs is that there may be differences in age groups, economic status, disability status, or any number of characteristics that might explain differences in outcomes instead of the program itself. To address this drawback, organizations might consider selecting populations that maximize similarity on basic demographic characteristics or employing an experimental design. Notwithstanding, whether one chooses to use experimental or implicit design depends on the program, the population of people it may affect, and the resources the organization has (Grembowski, 2016).

Cost-Benefit Analysis and Social Return on Investment

Cost-Benefit Analysis (CBA) and Social Return on Investment (SROI) are popular approaches to measuring social impact. These approaches monetize the direct and indirect outcomes of NPO services. Social Return on Investment (SROI) is the ratio of an organization’s community benefits to its resources or investments. The SROI represents the socio-economic return for every monetary unit spent (Manetti, 2014). SROI combines economic measures with social impact measures such as *social surplus*, *social benefit*, and *social costs* (Cordes, 2017).

CBA evaluates what the effects of a community program are on *social surplus*. Social surplus is whether a given community initiative produces outcomes with *social benefits*

that equal or exceed *social costs*. *Social benefits* are goods or services that provide positive value to some stakeholders. Some examples of social benefit could be the positive value experienced by families who are homeless when an anti-homelessness program provides shelter in a residential neighbourhood. More social benefit is generated from this effort through cost savings because on a larger scale, providing housing for families experiencing homelessness can save on municipal costs such as substance abuse supports, or public sanitation/clean-up (Cordes, 2017). *Social costs* are features of the program or organization that reduce its overall social value. An example of a social cost could be if residents of the neighbourhood where the shelters are do not want the shelters there or are worried about how it might affect their neighbourhood (Cordes, 2017).

Other social impact measures that SROI is concerned with are *transfers*, *willingness to pay*, and *externalities* (Cordes, 2017). *Transfers* are any program outcomes that neither take from nor add to social value, but instead shift existing social value from one segment of society to others. For example, if shifting funding from an anti-homelessness program to a substance abuse support program would have the same effect on homelessness and mental health, as well as costs to society, then this shift in funding would be a transfer.

Willingness to pay is the price that an individual or company would pay for a social good. The higher the price, the stronger the indicator that the social service provides value to its beneficiaries. NPOs can evaluate willingness to pay via *revealed preferences* or *stated preferences* of what people would pay for. Revealed preferences are based on actual behaviour (for example, observing what people pay for what goods and services). Stated preferences are indicated by survey results (for example, what maximum price people say they would be willing to pay). In general, revealed preferences are preferred by economists over stated preferences because they are directly observed in “real-life” situations where people actually spend their money, rather than stated as a hypothetical answer to survey questions. However, measuring willingness to pay through revealed or stated preferences is often more practical for social service providers.

Externalities are negative or positive effects experienced by other stakeholders (not those targeted by the program or service) (Cordes, 2017). Externalities can give a broader picture of social value, as it looks at the effect of the program or policy on wider society rather than just the supplier and the receiver. For example, if a program promoting pollinator gardens in a neighbourhood might not only have a positive effect on people who live there, but also those who pass by the gardens. Externalities are analogous to indirect outcomes in that a service or program can have social value beyond what was intended.

Finally, a few other concepts in CBA and SROI are *discounting rates* and *social opportunity cost*. *Discounting rates* look at the value of NPO services as time passes, and whether this value changes with time (Cordes 2017). For example, questions about a discounting rate include: how much will a neighbourhood value pollinator gardens this

year if they are planted this year? How much will the neighbourhood value them if they were to be planted next year, relative to how much they would value them if planted this year? *Social opportunity cost* outlines how using scarce resources to pursue a particular activity provides value to communities, but at the cost of not using those resources to fund another program that could also provide social value. Social opportunity cost can help investigate why one program may be better than another, because there is always another “next best” option to fund (Cordes, 2017). Taken together, all of these key concepts can help to build a picture of the SROI a program or SE offers, within the framework of CBA.

Stakeholder Involvement

Stakeholder involvement in the social impact measurement process was discussed in the literature as an important component. Stakeholder involvement can help increase buy-in for social impact measurement by supporting an organization with a greater number of networks for data collection (Alijani and Karyotis, 2019). Specific to social impact, stakeholder involvement is especially important for organizations to obtain additional perspectives on what is important to measure, the opportunities that may arise from measuring social impact, and community members that will be affected by measurement (Mook, Chan, and Kershaw, 2015). Research has suggested that the best way to include stakeholders in social impact measurement is to incorporate their community networks in the development of data systems and investing in the social impact measurement process (Alijani and Karyotis, 2019).

Another way to involve stakeholders is to use more inclusive evaluation methods that are grounded in participants’ experiences (rather than, for instance, offering closed-ended and predetermined questions on a survey). For instance, the ArtsSmarts Inquiry Model, developed to increase evaluation capacity within a Canadian arts education organization, allows for a description of lived experiences of users and stakeholders of NPO services (Chahine, 2012). This model focuses on strengthening ties between program users and providers and providing opportunities to learn from each other (Chahine, 2012). This framework may be useful to NPOs who operate as networks and require a flexible approach to evaluation (Chahine, 2012). As a form of collaborative inquiry, there are three phases in this model: Planning, Checking, and Gathering and Interpreting Data. In a school setting, *Planning* allows users to brainstorm answers to their own arts-related questions; *Checking* allows users to describe how their projects align with program curricula; and *Gathering and Interpreting Data* reflects what communities and users have learned from the service.

Indicators and Indicator Selection

Measuring impact requires carefully identifying and selecting indicators. There are several types of indicators that NPOs can use related to the level of effect they’re attempting to measure (i.e., output, outcome, or mission/impact). *Output-focused indicators* are typically a quantitative measure of what you delivered (e.g., the number of people accessing a service, the number of goods donated, the number of homeless

people sheltered) (Moxham, 2013). *Outcome-focused indicators* are used to evaluate the change created by an organization or service. For example, an outcome-focused indicator for a math tutoring program may be the number of students achieving at least a 'B' in their math course. Finally, *mission-focused (or impact) indicators* are created by NPOs by developing criteria of success based on their social mission or intended ultimate impact. These indicators are often harder to measure and may require following up with participants to assess the impact of a program or service. For instance, an impact-focused indicator for an experiential STEM education program for high school students might be the number of program participants who end up in STEM careers.

There are a number of ways to identify strong vs. weak indicators. Mark Friedman described three characteristics of strong performance indicators that have influenced performance measurement worldwide: data power (i.e., indicators developed from high-quality and unbiased methods), proxy power (i.e., indicators refer to an important and meaningful impact), and communication power (i.e., the indicator is relevant to broad and diverse audiences) (Friedman, 2000). In addition to these three characteristics, indicators should have a clear relationship to the organization's goals, should be easy-to-understand, should inform planning and policy, and should have adequate data management processes (Lichiello and Turnock, 1999). Strong performance indicators can also use "SMART" targets: specific, measurable, achievable, realistic and time-bound (Info Entrepreneurs, 2009). The structure of an indicator should include the following: a measure, target, data source, and reporting frequency. The *measure* refers to how the organization plans on quantifying that indicator. The *target* metric must match the measure and time period in which the indicator is being assessed. There must be a defined *data source* so that there is transparency and consistency in how indicators are measured. The *reporting* frequency varies per indicator and can include annually or biannually (Enochson, n.d.).

Consensus-based approaches such as the Delphi survey and nominal group technique are effective ways to prioritize indicators, particularly in diverse stakeholder groups, which in the case of arts and cultural organizations include service users, funders, and the broader community. The first step to both approaches is to develop a comprehensive list of relevant indicators that the organization can use to measure social impact. The Delphi method conducts at least two rounds of anonymous surveys where respondents rate their level of importance for each indicator (McMillan et al., 2016). An agreement level of 80% or higher usually indicates consensus on the importance of an indicator. An agreement level of 20% or lower represents consensus that the indicator is not important. For indicators with an agreement level between 20% and 80%, they are modified and translated into a second survey. During both surveys, respondents provide feedback on the wording and framing of indicators, which are used by the research team to refine indicators between survey rounds.

While the Delphi method involves anonymous surveys, which can be conducted online, the nominal group technique requires more active participation of respondents (McMillan et al., 2014). This method provides respondents with one question that they consider, which in the case of indicator selection can be: which indicators are important

for measuring the social impact for this art and cultural organization? The nominal group technique involves the following stages

1. Introduction to the NGT Process
2. Silent Generation: Respondents develop ideas in response to the initial question independently for 5 to 10 minutes.
3. Round Robin: All respondents describe their ideas until all are exhausted.
4. Clarification: Respondents clarify the meaning and intent of different ideas.
5. Ranking: Respondents rank the importance of ideas. First, ideas are categorized based on their similarity. Then, the categories of ideas are prioritized through discussion and informal voting.

Barriers to Measuring Social Impact

We found three barriers to measuring social impact: organizational culture, resource constraints, and accountability and relationships. First, *organizational culture* can be a barrier to social impact measurement when it creates conflict among employees. Conflict can arise from a mismatch in management’s evaluation goals and employees’ activities. Second, *resource constraints* influence the NPO’s ability to manage and measure social impact due to limited capacity and budget. Finally, *accountability and relationships* with stakeholders and funders determine the extent to which the results of evaluation activities can improve the social impact of NPO services.

Table 2: Descriptions of three categories of barriers to measuring social impact

<p>Organizational culture</p>	<p>NPOs that conduct social auditing or accountability evaluations have noticed changes to their organizational culture. Studies show that evaluations performed from external stakeholders can introduce discomfort and hinder staff morale (Arvidson, 2014; Bourgeois et al., 2015; Savaya and Waysman, 2005). Internal discomfort or conflict between employees and supervisors may develop due to a disagreement on how or why to measure performance. The language or the intrusive nature of questionnaires that measure social impact can potentially strain client and organization relationships based on differences of opinion. A paper authored by the Canadian Housing and Renewal Association draws on their experiences from a workshop held at their 2015 Congress on Housing and Homelessness. The paper focuses on how and why measuring social impact investments is advantageous to organizational goals and shares perspectives from different internal and external roles to the organization, such as the clients, housing provider, the provincial government agency, and the financial institution (Ervick-Knote, 2015). Furthermore, social impact evaluation can be more effective, genuine, and rigorous when initiated internally rather than being mandated by authorities (Savaya and Waysman, 2005).</p>
-------------------------------	--

Resource constraints	Resource constraints influence a NPO’s ability to organize, manage, and measure social impact. There are several challenges in collecting and analyzing data such as staffing, time constraints, operationalizing immeasurable outcomes, and a lack of direction or leadership approach (Barraket, 2013; Patton, 2001; Savaya and Waysman, 2005). The failure to integrate social impact measurement with other activities may also hinder the organization’s ability to allocate resources to evaluation activities (Patton, 2001). Evaluations can provide data for decision-making, identify variables that can lead to attaining desired outcomes, and assess the costs and benefits of activities.
Accountability and Relationships	Oftentimes, there is a lack of consensus on the goals and outcomes of social impact measurement between funding bodies and NPOs (Alijani, 2019; Arvidson, 2014; Benjamin, 2012). One study described a fundamental difference between charity-related and social responsibility relationships (Benjamin, 2012). For example, one study found that funders and NPOs behave in ways that reflect their self-interest (Campbell, 2016). Furthermore, funders may be apprehensive to reward NPOs solely based on socially good intentions instead of objective measurable results. An article authored by Charity Village presents how NPOs can benefit from measuring their impact, defining their theory of change, and explaining how they measure their intended outcomes (Charity Village, 2014). Moreover, organizations need to measure their social value and need to track their progress towards making a difference in their community or with their clients.

Facilitators for Measuring Social Impact

We found three common facilitators for measuring social impact: community outreach, data, and funding. First, *community outreach* is a major facilitator for conducting evaluation because it strengthens NPO relationships with their community by receiving direct feedback from them about the quality and impact of their services. Second, *data* facilitators are discussed in the context of the accessibility, usefulness, and cost of publicly available data. Finally, *funding* is a vital facilitator for social impact measurement. NPOs require sufficient funding to build the infrastructure and training conducive for measuring social impact.

Table 3: Descriptions of three categories of facilitators for measuring social impact

Category	Description
----------	-------------

Community Outreach	NPOs use their position in their community to strengthen community relationships. These relationships provide essential feedback from the community regarding the impact, accessibility, and reach of their services (Chahine, 2012). Direct and open communication may reveal indirect outcomes of NPO services (Barraket, 2013; Mackinnon and Stephens, 2010; Molecke and Pinkse, 2017; Molecke and Pinkse, 2020). Identifying the resources in a community that a NPO can leverage to evaluate their program facilitates social impact measurement. For example, open houses, volunteers, and committees are effective methods for engaging with the community in evaluation (CallHub, 2020).
Data	Publicly available data may provide additional support for NPO social impact measurement activities. Publicly available data can be retrieved from health departments, education boards, census data, or private foundations (Minich, 2006; Pennerstofer and Rutherford, 2019).
Funding	The ability to measure social impact requires significant funding. Funding organizations recognize that creating an effective plan to measure social impact is often time and cost intensive. Measuring social impact requires a team of researchers who collect, manage, analyze, and report data. The major costs are associated with data collection and management. Funders often require organizations to identify all of their major costs associated with measuring social impact which shows funders that the organization has carefully considered expenses and opportunities.

Recommendations

Recommendation 1: Determine the purpose of social impact measurement

A vision is a statement that clarifies the overall purpose of evaluation to employees and other stakeholders (Latham, 1995). A well-articulated vision statement for social impact measurement may enable employees to better understand their roles and responsibilities with respect to evaluation activities (Mitchell, 2016). We have found that NPOs adopt several purposes for social impact measurement: (1) as a requirement for funding, (2) to improve the quality, reach, and impact of services, (3) to increase managerial control over employee activities, and (4) to market their services to stakeholders and the broader community (Liket et al., 2014). We recommend that the DVSA’s first step in evaluating their social impact is to develop a vision and purpose of evaluation that might emphasize improving the quality of their services for the communities they serve. The second step is to communicate this vision to service users, broad community, staff, and funding agencies. We believe that having a robust vision that primarily focuses on service quality will improve the quality and usefulness of social impact measurement, increase the likelihood of expanding existing services, and acquiring new funding.

Recommendation 2: Develop a robust social impact measurement framework informed by lessons from diverse disciplines and context, but that is unique to the needs of the DVSA

This report provides a description of a selection of frameworks and approaches to measuring social impact: logic models, cost benefit analysis, and social return on investment. The DVSA can use this information to develop a tailored social impact measurement system specific to their context. Since these frameworks come from distinct disciplines, they emphasize different components of social impact. The DVSA will be able to consult these descriptions to identify which elements of social impact measurement are most important to their organization and the communities they serve.

Recommendation 3: Consider including health indicators to measure social impact

We recommend that the DVSA consider including health indicators in their social impact evaluation framework. Some indicators that the DVSA may want to look at include users' self-reported ratings of their overall health, quality of life, confidence, or happiness before and after participating in DVSA programs. Some helpful tools which can measure these include surveys developed by researchers and international organizations. One example is the *SF-36*, a survey developed by RAND Health Care, which asks questions about overall physical and emotional health (RAND, 2020). The DVSA may also consider using the *Social and Emotional Loneliness Scale for Adults* (DiTomasso and Spinner, 1993), or the *UCLA Loneliness Scale* (Russel et al., 1978; Fetzer Institute, n.d.) to survey their participants and learn about any feelings of loneliness they may have, and how those feelings might change before and after program participation. If feasible, the DVSA may consider hiring trained professionals or clinically trained volunteers to measure physical markers of health, such as blood pressure or memory and orientation, before and after art therapy participation. These may be useful measures in older adult participants living in long-term care, who may be at risk for high blood pressure or memory issues and experience physical health benefits as a result of program participation. By emphasizing the impact that their work has on health, the DVSA will be better able to compete for a vast selection of health funding and catalyst grants to continue their important work. Including health indicators might expand the pool of funding the DVSA is eligible for with only minor changes to existing services and activities.

Recommendation 4: Involve stakeholders (users, funders, and art professionals) in the development and implementation of social impact measurement

Stakeholder involvement in developing social impact measurement methods is essential because it increases the effectiveness, perceived usefulness and uptake of social

impact measurement. For example, Liket et al. (2014) found that NPOs were much more likely to judge their evaluations as useful and effective when they were designed in collaboration with stakeholders. Stakeholders in the arts include more than just the funders; they include customers, consumers, community, and art professionals, all of whom should advise on the scope of their social impact evaluation and the processes used. It is important to note that social impact measurement is not stable. The needs of communities will shift over time and therefore measuring social impact may need to adapt to community needs. Ongoing stakeholder consultation and involvement may enable the DVSA to ensure that their methods and indicators for measuring social impact remain responsive to their community. The DVSA can conduct one survey and a set of semi-structured interviews annually to identify community needs, preferences, and expectations, and ensure that their social impact measurement aligns with the communities they serve.

Recommendation 5: Strengthen organizational capacity for high-quality social impact measurement

Building capacity for social impact measurement is challenging because it requires a considerable amount of financial and non-financial resources. To build capacity for social impact measurement, the DVSA must consider how they plan to contribute to a larger organization culture (Campbell, 2016; Morales, 2018; Wing, 2004). Since we have found that commitment and cohesion within the organization to measure social impact is an important facilitator (Lynch-Cerullo and Cooney, 2011; Minich, 2006), the DVSA's leadership team may consider continuously expressing the importance of social impact measurement to staff and the community they serve.

In this report, we described a number of barriers and facilitators that the DVSA might face when attempting to measure social impact. While these barriers were identified by analyzing the experiences of other NPOs, we believe that an organizational culture that values evaluation and managerial commitment to social impact measurement are the most important to consider given the current highly competitive funding climate for arts and cultural organizations. Identifying the internal and external barriers to social impact measurement may allow the DVSA to plan for possible obstacles. The DVSA should strengthen their evaluation culture through a number of strategies that include developing a clear vision, communicating to employees regularly about the purpose and impact of evaluation through workshops, and ensuring that management understands the mission and purpose of social impact measurement.

Conclusion

This report described the findings of a database review and grey literature search of frameworks, indicators, barriers, facilitators, and resources for measuring performance and social impact in NPOs. We described a number of frameworks that the DVSA can use to inform the development of their own framework such as logic models and Social

Return on Investment. We also offer important considerations when selecting the appropriate indicators to measure performance. We offer five recommendations as next steps for the DVSA (1) determine the purpose of social impact measurement; (2) develop a robust social impact measurement framework informed by lessons from diverse disciplines but that is unique to the needs of the DVSA; (3) consider including health indicators to measure social impact (4) involve stakeholders (users, funders, and art professionals) in the development and implementation of performance measurement; and (5) strengthen organizational capacity for high-quality performance measurement.

Bibliography

Alijani, S., & Karyotis, C. (2019). Coping with impact investing antagonistic objectives: A multistakeholder approach. *Research In International Business And Finance*, 47, 10-17. <https://doi.org/10.1016/j.ribaf.2018.04.002>

Arena, M., Azzone, G., & Bengo, I. (2015). Performance Measurement for Social Enterprises. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, 26(2), 649–672. <https://doi.org/10.1007/s11266-013-9436-8>

Arvidson, M., & Lyon, F. (2013). Social Impact Measurement and Non-profit Organisations: Compliance, Resistance, and Promotion. *VOLUNTAS: International Journal Of Voluntary And Nonprofit Organizations*, 25(4), 869-886. <https://doi.org/10.1007/s11266-013-9373-6>

Barraket, J., & Yousefpour, N. (2013). Evaluation and Social Impact Measurement Amongst Small to Medium Social Enterprises: Process, Purpose and Value. *Australian Journal Of Public Administration*, 72(4), 447-458. <https://doi.org/10.1111/1467-8500.12042>

Bourgeois, I., Whynot, J., & Thériault, É. (2015). Application of an organizational evaluation capacity self-assessment instrument to different organizations: similarities and lessons learned. *Evaluation and program planning*, 50, 47–55. <https://doi.org/10.1016/j.evalprogplan.2015.01.004>

Benjamin, L. M. (2012). Nonprofit Organizations and Outcome Measurement: From Tracking Program Activities to Focusing on Frontline Work. *American Journal of Evaluation*, 33(3), 431–447. <https://doi.org/10.1177/1098214012440496>

CallHub. (2020). Guide to Devising Great Community Outreach Strategies That Win Big. (2020, June 10). Retrieved from <https://callhub.io/community-outreach-strategies/>

Campbell, D., & Lambright, K. (2016). Struggling to Get It Right. *Nonprofit Management And Leadership*, 27(3), 335-351. <https://doi.org/10.1002/nml.21249>

Chahine, S. (2012). Networked Based Evaluation: ArtsSmarts Inquiry Model. *Systemic Practice and Action Research*, 25(5), 407-416.

Charity Village. (2014). Measuring Impact: How small and medium-sized nonprofits can benefit from effectively measuring their impact. Charity Village.
<https://charityvillage.com/measuring-impact-how-small-and-medium-sized-nonprofits-can-benefit-from-effectively-measuring-their-impact/>

Coalter, F. (2001). *Cultural Services: Realising the Potential*, Research Agenda. University of Edinburgh: Centre for Leisure Research.

Cordes, J. J. (2017). Using cost-benefit analysis and social return on investment to evaluate the impact of social enterprise: Promises, implementation, and limitations. *Evaluation and Program Planning*, 64, 98–104.
<https://doi.org/10.1016/j.evalprogplan.2016.11.008>

DiTommaso, E., & Spinner, B. (1993). The development and initial validation of the Social and Emotional Loneliness Scale for Adults (SELSA). *Personality and Individual Differences*, 14(1), 127–134. [https://doi.org/10.1016/0191-8869\(93\)90182-3](https://doi.org/10.1016/0191-8869(93)90182-3)

Enochson, H. (n.d.). 27 Examples of key performance indicators. Retrieved from <https://onstrategyhq.com/resources/27-examples-of-key-performance-indicators/#:~:text=The%20best%20KPIs%20have%20more,time%20period%20of%20your%20goal.&text=Reporting%20Frequency%20%E2%80%93%20Different%20KPIs%20may,on%20them%20at%20least%20monthly.>

Ervick-Knote, H. (2015). Social Impact Measurement: Making the Case for Housing. *The CHRA Congress Sessions Series 2015*. BC Housing. <https://chra-achru.ca/wp-content/uploads/2015/08/Social-Impact-Measurement.pdf>

Fetzer Institute. (n.d.). UCLA Loneliness Scale. Retrieved August 23, 2020, from https://fetzer.org/sites/default/files/images/stories/pdf/selfmeasures/Self_Measures_for_Loneliness_and_Interpersonal_Problems_UCLA_LONELINESS.pdf

Friedman, M. (2000). Results Accountability for Proposition 10 Commissions: A Planning Guide for Improving the Well-Being of Young Children and Their Families. Building Community Systems for Young Children.

Grembowski, D. (2016). *The practice of health program evaluation* (2nd ed.). SAGE Publications, Inc.

Hager, M. A. (2001). Financial Vulnerability among Arts Organizations: A Test of the Tuckman-Chang Measures. *Nonprofit and Voluntary Sector Quarterly*, 30(2), 376–392.
<https://doi.org/10.1177/0899764001302010>

Harbour, J. L. (2017). *The basics of performance measurement*. Crc Press.

Ibrisevic, I. (2018). Measuring Nonprofit Social Impact: A Crash Course. Donorbox Nonprofit Blog. <https://donorbox.org/nonprofit-blog/measuring-nonprofit-social-impact/>

Info Entrepreneurs. (2009). Measure performance and set targets. Retrieved from <https://www.infoentrepreneurs.org/en/guides/measure-performance-and-set-targets/#:~:text=It's%20a%20familiar%20acronym%2C%20but,definition%2C%20be%20specific%20and%20measurable.>

Jermyn, H. (2001). The Arts and Social Exclusion: A review prepared for the Arts Council of England. The Arts Council of England. <https://culturelab.net/public/articles/c1c671fd26ef8e10fbc1f7e5edd553247ff9e7ba90fc6a4bc89425a493c8/download/3764>

Kishchuk, P. (2014, December 10). Evaluation of the School of Visual Arts: Findings synthesis. Yukon.ca. <https://yukon.ca/en/evaluation-school-visual-arts-2014>

Lall, S. A. (2019). From Legitimacy to Learning: How Impact Measurement Perceptions and Practices Evolve in Social Enterprise–Social Finance Organization Relationships. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, 30(3), 562–577. <https://doi.org/10.1007/s11266-018-00081-5>

Landry, C., Bianchini, F., Maguire, M., & Worpole, K. (1993). The Social Impact of the Arts A Discussion Document. *Comedia*.

Latham, J. R. (1995). Visioning: The concept, trilogy, and process. *Quality Progress*, 28(4), 65-70.

Lee, C., & Nowell, B. (2015). A Framework for Assessing the Performance of Nonprofit Organizations. *American Journal of Evaluation*, 36(3), 299–319. <https://doi.org/10.1177/1098214014545828>

Lichiello, P., & Turnock, B. J. (1999). *Guidebook for Performance measurement*. Turning Point.

Liket, K. C., Rey-Garcia, M., & Maas, K. E. H. (2014). Why Aren't Evaluations Working and What to Do About It: A Framework for Negotiating Meaningful Evaluation in Nonprofits. *American Journal of Evaluation*, 35(2), 171–188. <https://doi.org/10.1177/1098214013517736>

Lynch-Cerullo, K., & Cooney, K. (2011). Moving from outputs to outcomes: A review of the evolution of performance measurement in the human service nonprofit sector. *Administration in Social Work*, 35(4), 364-388.

Mackinnon, S., & Stephens, S. (2010). Is participation having an impact? Measuring progress in Winnipeg's inner city through the voices of community-based program participants. *Journal of Social Work, 10*(3), 283-300.

Manetti, G. (2014). The Role of Blended Value Accounting in the Evaluation of Socio-Economic Impact of Social Enterprises. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations, 25*(2), 443–464. <https://doi.org/10.1007/s11266-012-9346-1>

McMillan SS, King M, Tully MP. How to use the nominal group and Delphi techniques. *International journal of clinical pharmacy. 2016 Jun 1;38*(3):655-62.

McMillan, S. S., Kelly, F., Sav, A., Kendall, E., King, M. A., Whitty, J. A., & Wheeler, A. J. (2014). Using the nominal group technique: how to analyse across multiple groups. *Health Services and Outcomes Research Methodology, 14*(3), 92-108.

Mitchell, G. E., & Berlan, D. (2016). Evaluation and evaluative rigor in the nonprofit sector. *Nonprofit Management and Leadership, 27*(2), 237-250.

Minich, L., Howe, S., Langmeyer, D., & Corcoran, K. (2006). Can community change be measured for an outcomes-based initiative? A comparative case study of the success by 60 Initiative. *American journal of community psychology, 38*(3-4), 183-190.

Molecke, G., & Pinkse, J. (2017). Accountability for social impact: A bricolage perspective on impact measurement in social enterprises. *Journal of Business Venturing, 32*(5), 550-568.

Molecke, G., & Pinkse, J. (2020). Justifying Social Impact as a Form of Impression Management: Legitimacy Judgements of Social Enterprises' Impact Accounts. *British Journal of Management, 31*(2), 387-402.

Mook, L., Chan, A., & Kershaw, D. (2015). Measuring social enterprise Value creation: The case of furniture bank. *Nonprofit Management and Leadership, 26*(2), 189-207.

Morales, J. R., Ferron, C., Whitmore, C., Reifel, N., Geary, E., Anderson, C., & Mcdaniel, J. (2018). Performance measurement in tribal home visiting: Challenges and opportunities. *Infant mental health journal, 39*(3), 312-325.

Moxham, C. (2013). Measuring up: examining the potential for voluntary sector performance measurement to improve public service delivery. *Public Money and Management, 33*(3), 193-200. <https://doi.org/10.1080/09540962.2013.785704>

Ontario Arts Council. (n.d.). Performance measurement framework. [https://www.arts.on.ca/oac/media/oac/Publications/Research%20Reports%20EN-FR/Performance%20Measures/OAC Performance Measures Framework EN October 2019 FINAL Accessible.pdf](https://www.arts.on.ca/oac/media/oac/Publications/Research%20Reports%20EN-FR/Performance%20Measures/OAC%20Performance%20Measures%20Framework%20EN%20October2019%20FINAL%20Accessible.pdf)

Ontario Trillium Foundation. (2020). What We Fund. (n.d.). Retrieved from <https://otf.ca/what-we-fund?redirected=1>

Patton M. Q. (2001). The CEFPP as a model for integrating evaluation within organizations. *Cancer practice*, 9(Suppl 1), S11–S16. <https://doi.org/10.1046/j.1523-5394.2001.95103.x>

Pennerstorfer, A., & Rutherford, A. C. (2019). Measuring growth of the nonprofit sector: The choice of indicator matters. *Nonprofit and Voluntary Sector Quarterly*, 48(2), 440-456.

Polonsky, M., & Grau, S. L. (2011). Assessing the social impact of charitable organizations—four alternative approaches. *International Journal of Nonprofit and Voluntary Sector Marketing*, 16(2), 195–211. <https://doi.org/10.1002/nvsm.407>

RAND (2020). *36-Item Short Form Survey from the RAND Medical Outcomes Study*. Retrieved August 23, 2020, from https://www.rand.org/health-care/surveys_tools/mos/36-item-short-form.html

Reid, W., & Karambayya, R. (2009). Impact of dual executive leadership dynamics in creative organizations. *Human relations*, 62(7), 1073-1112.

Russell, D., Peplau, L.A., & Ferguson, M.L. (1978) Developing a measure of loneliness. *Journal of Personality Assessment*, 42, 290-294.

Savaya, R., & Waysman, M. (2005). The Logic Model. *Administration In Social Work*, 29(2), 85-103. https://doi.org/10.1300/j147v29n02_06

Turbide, J., & Laurin, C. (2009). Performance measurement in the arts sector: The case of the performing arts. *International journal of arts management*, 56-70.

Wing, K. T. (2004). Assessing the effectiveness of capacity-building initiatives: Seven issues for the field. *Nonprofit and Voluntary Sector Quarterly*, 33(1), 153-160.

W.K. Kellogg Foundation. (2004, Jan). *W.K. Kellogg Foundation Logic Model Development Guide*. Retrieved from <https://www.bttop.org/sites/default/files/public/W.K.%20Kellogg%20LogicModel.pdf>

Appendices

Appendix A: Search Strategy

Social Sciences Citation Index (SSCI) + Arts and Humanities Citation Index (A&HCI) via Web of Science (2000-2020) (n = 527)

1/ TS=(nonprofit\$ OR non-profit\$ or nonprofit organization\$) OR TI=(nonprofit\$ OR non-profit\$ or nonprofit organization\$ OR social enterprise* OR social purpose organization\$)

2/TS=(cultur* OR art* OR leisure* OR recreation* OR communit* OR social*) OR TI=(cultur* OR art* OR leisure* OR recreation* OR communit* OR social*)

3/ TI=(impact\$ OR measur* OR assess* OR evaluat* OR metric\$ OR indicator\$ OR impact model\$ OR impact framework\$ OR assessment model\$ OR assessment framework\$ OR measurement model\$ OR measurement framework\$ OR evaluation model\$ OR evaluation framework\$ OR causal model\$ OR causal framework\$)

Appendix B: Descriptive Characteristics of Academic Literature

Author, year	Title of study	Research objectives	Country	Study Design (Primary, Review, Case Study, or Commentary)	Abstract Findings
Alijani, 2019	Coping with impact investing antagonistic objectives: A multistakeholder approach	To describe impact investing and the web of interactions among numerous stakeholders with opposing objectives and antagonistic assets in sourcing and channeling financial resources	Global	Review	This paper highlights the importance of adopting a multistakeholder approach when examining the financial ecosystem within which impact investing intermediaries operate and collaborate to achieve maximum social and environmental impact.
Arena, 2015	Performance measurement for social enterprises	Develop an approach that could be applied to by socioeconomic status to measure their results with respect to social, environmental, and economic impacts	USA	Review	The proposed approach consists of a “general” PMS model for SEs—i.e., the performance dimensions that should be measured—and a stepwise method to be used by SEs to develop their own PMS. For sake of clarification, the proposed approach is applied to the case of an Italian SE competing in the energy sector to develop a set of key performance indicators.
Arvidson, 2014	Social Impact Measurement and Non-profit	An examination of the experience and behaviour of non-profit	UK	Primary	While most organizations can relinquish control when tasked with evaluation, they can also resist

	Organizations: Compliance, Resistance, and Promotion.	organizations in the UK in relation to a demand for social impact evaluations			through using discretion in deciding what and how to measure and report. Social impact measurement can also be used for learning and promotional purposes, and as a way of exerting control over their environment. Strategic decoupling is used to explain the differences observed between what organizations are asked to do, what they plan to do, and what they actually do.
Bagnoli, 2011	Measuring performance in social enterprises	Addresses the subject of management control and creates a performance measurement system for social enterprises	Italy	Review	Three reference fields for management are analyzed: (a) economic-financial performance, linked to the determination of general performance (profits, value added, etc.) and analytic results (production-cost of services, efficiency indicators, etc.); (b) social effectiveness, to measure the quantity and quality of work undertaken and to identify its impact on the intended beneficiaries and the community; and (c) institutional legitimacy, verifying conformity with law and mission statement. In integrating these aspects, they propose a multidimensional controlling framework that is appropriate to the management of a social enterprise. They try to demonstrate the validity of the theoretical model through the

					illustration of an Italian empirical case: the work integration social enterprise "Ulisse."
Banning, 2007	Maximizing public relations with the organization-public relationship scale: Measuring a public's perception of an art museum	To measure member perceptions of an art museum using an organization-public relationship scale	USA	Primary	Analysis shows member perceptions of the museum-public relationship differentiated members likely to continue their membership from those likely to discontinue their membership with the museum. This study confirms the appropriateness of using the organization-public relationship scale with museums.
Barraket, 2013	Evaluation and Social Impact Measurement Amongst Small to Medium Social Enterprises: Process, Purpose and Value	To examine the practices and effects of evaluation and impact measurement by five small to medium social enterprises in Australia	Australia	Primary/Review	The findings from this research suggest that, at the organizational level, the primary perceived benefit of undertaking evaluation and impact measurement by small to medium social enterprises is to advance organizational learning and, thus, performance. However, in practice, the dominant driver of completing these activities amongst our participants was to demonstrate legitimacy to external stakeholders.
Bartual Sanfeliu et al., 2013	Measuring performance of social and non-profit Microfinance Institutions	Proposes a multi-criterion methodology based on goal programming that simultaneously considers different	Latin America	Primary	NR

	(MFIs): An application of multi-criterion methodology	categories involved in the performance of MFIs			
Benjamin, 2010	Mediating accountability	Non-profit funding intermediaries and use of performance measurement	USA	Primary	NR
Benjamin, 2010	Funders as Principals Performance Measurement in Philanthropic Relationships	Examination/description of tension that performance measurement brings up in philanthropy -- philanthropic relationship are not equivalent to accountability relationships.	USA	Review	Due to the differences in philanthropic and accountability relationships, without a systematic way to attend to concerns when adopting performance measurement, funders must choose to their reward measurable results or fall back on supporting good intentions
Benjamin, 2012	The Potential of Outcome Measurement for Strengthening Nonprofits' Accountability to Beneficiaries	To examine the guidance given to non-profits about how to develop an outcome measurement system for their program or organization, and to understand how and in what ways outcome measurement could strengthen downward accountability and then analyzes the outcome measurement guides	USA	Primary	The analysis shows that the guides were neither uniform in the conceptualization of non-profit beneficiaries nor in how they directed non-profits to use outcome measurement with their beneficiaries. Despite scholars' suggestion that a non-profit's relationship to their beneficiaries is a key accountability relationship, the guides suggest that beneficiaries have an ambiguous standing, relative to other stakeholders, in the non-profit accountability environment.

		targeted to non-profits along these dimensions			
Benjamin, 2012	Nonprofit Organizations and Outcome Measurement: From Tracking Program Activities to Focusing on Frontline Work	Inform better efforts of impact measurement in non-profits and examine the extent that existing frameworks actually measure what non-profit staff do to support positive outcomes for their community.	USA	Review/Commentary	Existing outcome measurements only focus on how staff implement programs instead of how they work with the community.
Boorsma, 2010	Arts Marketing Performance: An Artistic-Mission-Led Approach to Evaluation	To describe evaluating the performance of arts marketing and propose a model to guide this evaluation	Netherlands	Primary/Review	This article suggests evaluating the performance of arts marketing based on the contribution made to the achievement of the arts organization's artistic mission and proposes a model based on Kaplan and Norton's Balanced Scorecard to guide the artistic-mission-led evaluation of arts marketing performance.
Bourgeois et al., 2015	Application of an organizational evaluation capacity self-assessment instrument to different organizations: Similarities and	Contribute to growing body of knowledge; summary of an early attempt to measure EC in three different organizations; apply the organizational evaluation capacity self-assessment instrument to varying	Canada	Primary	Evaluation capacity tends to be higher, both in terms of capacity to do and capacity to use, in organizations that have developed systematic mechanisms to institute an evaluation culture; capacity does not first require capacity to do

	lessons learned	organizational contexts; identify the organizational characteristics that are specific to each organizational type and that may have an impact on evaluation capacity.			
Brook & Akin, 2019	Using Theory of Change as a Framework for Examining Community Context and Philanthropic Impact	Examine evaluation approaches used by grantees/assess a foundation's theory of change by identifying topics of interest and coding grantee proposals if find topics.	USA	Primary	Community agencies were improving in their uptake of evidence-based interventions and use of validated instruments. Grantees needed support in building agency capacity
Brooks, 2004	Evaluating the effectiveness of nonprofit fundraising	Describes and compares common quantitative measures of fundraising performance	USA	Primary	NR
Calo et al., 2019	Exploring the Contribution of Social Enterprise to Health and Social Care: A Realist Evaluation	Uncover how provisions by social enterprises might achieve positive health outcomes	UK	Primary	Social enterprises are better able to flexibly deliver a programme designed around the needs of service users. Their role as a community "boundary spanner" help facilitate strong ties and feelings of connectedness--and funding has a strong moderating effect.
Campbell, 2016	Struggling to Get it Right:	Description of common issues that public and	USA	Primary	Some common issues include dissatisfactions with formal data

	Performance Measurement Challenges and Strategies for Addressing them among funders of Human Services	non-profit FUNDERS face - from their perspective, rather than the organization/service providers'			collection procedures, difficulty getting providers to comply with reporting requirements, provider performance problems, lack of capacity to use performance information
Cannatelli, 2017	Exploring the Contingencies of Scaling Social Impact: A Replication and Extension of the SCALERS Model	To provide theoretical grounding through contingency theory and to conduct the first empirical test of the situational contingencies of the SCALERS model	Italy	Primary/Review	A positive relationship between each of the SCALERS variables and scaling—except replicating—has been found. Initial evidence of five contingencies that moderate the relationship between the SCALERS and scaling of social impact has also been found.
Carman, 2010	Evaluation capacity and nonprofit organizations: is the glass half-empty or half-full	Explore the evaluation capacity for today's non-profit organizations	USA	Primary	There are three types of non-profit organizations. The first type of non-profit organization is one that, by most accounts, is satisfied with their evaluation efforts. Although these organizations report that they struggle with not having as much time as they would like to devote to evaluation, they are fairly satisfied with their levels of evaluation expertise and report having few problems with the implementation of evaluation systems. The second type of organization has some struggles

					with evaluation. These organizations report having internal support for evaluation from management, the board and staff, and some capacity to implement an evaluation system, yet they struggle with evaluation design issues, data collection, and resources for evaluation. The third type of organization is one that is struggling across the board. These organizations report having substantial implementation challenges, in terms of lacking basic resources (i.e., staff, funding, time), lacking evaluation expertise, and they report having little support for evaluation from funders, the board, management, and staff. We conclude by exploring the implications of these findings.
Carman, 2009	Nonprofits, Funders, and Evaluation - Accountability in Action	to examine the extent different types of funders are asking non-profits for evaluation and performance measurement data.	USA	Primary	Non-profits that receive considerable funding from the federal government or United Way are engaging in program evaluation and performance measurement.
Chahine, 2012	Networked Based Evaluation: ArtsSmarts Inquiry Model	To describe the ArtsSmarts Inquiry Model (AIM) which allows networks of partners to conduct program evaluations	Canada	Commentary	NR

Cohen et al., 2003	Arts and economic prosperity: the economic impact of nonprofit arts organizations and their audiences	Highlight methods and findings on the economic impact of non-profit arts organizations and their audiences	USA	Primary	NR
Cordes, 2017	Using cost-benefit analysis and social return on investment to evaluate the impact of social enterprise: Promises, Implementation, and limitations	Description/overview of CBA and SROI as tools for setting funding priorities and evaluating performances	USA	Commentary	There are some limitations/issues raised in the implementation of CBA and SROI, yet they can be a useful lens for setting priorities and/or evaluating performance
Coryn et al., 2009	Adding a Time-Series Design Element to the Success Case Method to Improve Methodological Rigor - An Application for Nonprofit	To modify the Success Case Method (SCM) to define success in a context that did not have an observable, measurable ROI or impact, and to add a longitudinal, time-series design element	USA	Commentary	NR

	Program Evaluation				
Crucke, 2016	The Development of a Measurement Instrument for the Organizational Performance of Social Enterprises	To develop an instrument suitable for the internal assessment and the external reporting of the non-financial performance of a diverse group of social enterprises	Belgium	Primary/Review	Focusing on five dimensions of organizational performance (economic, environmental, community, human and governance performance), the study offered a set of indicators and assessment tool for social enterprises.
Davis, 2020	Dancing in the Street: Impacting At-Risk Youths' Lives through the Arts	Destiny Arts Youth Performance Company - a non-profit, community dance group based in Oakland CA, offers competitive scholarships to deserving at-risk youth in lower class, largely ethnic communities. examination of an ethnographic documentary - looked at the lives of five youths whose lives are affected and transformed through the creative medium of dance	USA	Review	For at-risk and diverse teens, collaborative art can be a foundation for personal strength, liberation, and hope.

Despard, 2016	Strengthening evaluation in nonprofit human service organizations: results of a capacity-building experiment	Review NPHSO capacity-building interventions	NR	Primary	Non-profit human service organizations that receive evaluation-related assistance experience gains
Eckerd & Moulton, 2011	Heterogeneous Roles and heterogeneous practices: understanding the adoption and use of nonprofit performance evaluations	Assess non-profit evaluation practices	NR	Primary	Institutional perceptions are consistent predictors of whether non-profit organizations adopt particular practices
Eller, 2018	Nonprofit Organizations and Community Disaster Recovery: Assessing the Value and Impact of Intersector Collaboration	To investigate the value and impact of those disaster recovery services provided by the non-profit sector under an intersector operational collaborative model between voluntary NPOs and state and federal agencies	USA	Review	The study demonstrates that the beneficial impact of non-profit organizations on community recovery is more far-reaching than is typically measured. Further, the analysis also sheds light on those key factors driving successful collaboration between the public and non-profit sectors in disaster management. Beyond the particulars of the case, this assessment identifies the broader implications of this general type of

					management strategy for future disaster recovery efforts.
Embrahim, 2014	What Impact? A Framework for Measuring the Scale and Scope of Social Performance	To build a performance assessment framework premised on an organization's operational mission, scale and scope	USA	Primary/Review	Not all organizations should measure their long-term impact; rather, some organizations would be better off measuring shorter-term outputs or individual outcomes.
Fowler et al., 2017	Transforming Good Intentions into Social Impact: A Case on the Creation and Evolution of a Social Enterprise	Tries to show the application of a process model about sustaining a social enterprise with a mission. Also explores the social value creation process.	USA	Case study	social enterprises use operating models to show social impact, but these models and their resources continue to evolve
Gaber, 2000	Meta-Needs Assessment	Examine the use of meta-needs assessment in organizations	USA	Review	NR
Gazley & Abner, 2014	Evaluating a Product Donation Program: Challenges for Charitable Capacity	To describe the impact of a national product donation program (includes numbers served, a faith-based mission, ability to process product donations effectively, continued participation)	USA	Primary	Experience with product donation programs partly mitigates the challenges of participation and that success may be due to how central the program is to organizational objectives

Gooch et al., 2015	Impact in Interdisciplinary and Cross-Sector Research: Opportunities and Challenges	To investigate the practices of interdisciplinary projects and see how they create impact	UK	Primary	Interdisciplinary projects offer a unique perspective that in turn also encourage a lot of sustainable outcomes. A lack of accountable designs and collaboration practices may hinder pathways to impact
Gordon et al., 2017	How Can Social Enterprises Impact Health and Well-Being?	To examine the impacts of social enterprise on individual and community health and well-being (determinants of health and the influence of structure on their outcomes)	UK	Case study	Social enterprises impacted all layers of the social determinants of health; however, the greater impact was on individual lifestyle factors and social and community networks. Structural factors can enable and constrain impact on all levels.
Grieco, 2015	Measuring Value Creation in Social Enterprises: A Cluster Analysis of Social Impact Assessment Models	To develop a hierarchical cluster analysis based on a sample of 76 SIA models and group them in macro-categories in order to help social entrepreneurs choose the model that is best suited to the needs of their organization	Italy	Primary/Review	In this research, a hierarchical cluster analysis was developed based on a sample of 76 SIA models to group them in macro-categories and help social entrepreneurs choose the model that is best suited to the needs of their organization.
Hager, 2001	Financial Vulnerability Among Arts Organizations:	an empirical test of Tuckman-Chang measures of financial vulnerability to predict	USA	Primary	Although Tuckman-Chang measures do not have utility for all types of arts non-profits, the measures still show substantial promise for predicting the

	A Test of the Tuckman-Chang Measures	financial vulnerability of arts-based non-profits			closure of some non-profit organizations. NOTE: due to study design they were limited to large arts organizations which are functionally and financially different than small ones
Jones & Mucha, 2013	Sustainability Assessment and Reporting for Nonprofit Organizations: Accountability "for the Public Good"	To assess sustainability reporting frameworks and discuss the challenges of implementing these programs	USA	Case study	NR
Kim & Ryzin, 2014	Impact of Government funding on donations to Arts organizations: A survey experiment	Examine the behavioural aspect of the crowding-out hypothesis	NR	Primary	Non-profit with government funding half likely to receive all money in a forced-choice scenario
Kluger, 2006	The Program Evaluation Grid: A Planning and Assessment Tool for Nonprofit Organizations	to give an overview of the Program Evaluation Grid tool, which helps organizations to rank the effectiveness about their programs, expand those which are highly ranked, and improve those ranked low	USA	Review	the tool can be used to generate a rank which is useful for key decision-making processes such as budgeting and program reduction or elimination

Krawczyk, Wooddell, & Dias, 2017	Charitable Giving in Arts and Culture Nonprofits: The Impact of Organizational Characteristics	Intend to better answer questions regarding the impact of specific charitable giving, and whether organizational characteristics affect various categories of donations	USA	Primary	"Allows for more informed fundraising decisions"
Lall, 2017	Measuring to Improve Versus Measuring to Prove: Understanding the Adoption of Social Performance Measurement Practices in Nascent Social Enterprises	To better define social performance measurement	USA	Primary	Adopting social performance measurement practices in social enterprises is related to the growing rationalization of the social sector
Lall, 2019	From Legitimacy to Learning: How Impact Measurement Perceptions and Practices Evolve in Social Enterprise-Social Finance	an examination of how social enterprises interact with social finance organizations in the context of impact measurement	Global	Primary	Relationships between enterprise and funder are hierarchical and rigid at early stages, but over time evolve into more collaborative partnerships. Eventually, SEs embrace impact measurement as a tool for organizational learning, and social finance organizations develop more empowering approaches for impact measurement

	Organization Relationship				
Leak et al., 2007	Evaluating the capacity of faith-based programs in Colorado	Examine faith-based organizations' capacity gains after participating in a targeted capacity-building intervention	USA	Primary	Targeted technical assistance predicted improvements in organizational capacity; faith-based organizations showed significant improvements in almost all areas of capacity; comparisons between faith- and community-based organizations showed few differences in capacity gains
Lee, 2015	A framework for assessing the performance of nonprofit organizations	Review contemporary models of non-profit performance measurement to develop an integrated framework in order to identify directions for advancing the study of performance measurement	USA	Review	Our analysis of this literature illuminates seven focal perspectives on non-profit performance, each associated with a different tradition in performance measurement. Second, we demonstrate the utility of this integrated framework for advancing theory and scholarship by leveraging these seven perspectives to develop testable propositions aimed at explaining variation across non-profits in the adoption of different measurement approaches. By better understanding how performance measurement is conceptualized within sector, the field will be better positioned to both critiques and expand upon normative approaches advanced in the literature as well as advance theory for predicting performance measurement decisions.

Liket, 2014	Why aren't evaluations working and what to about it: a framework for negotiating meaningful evaluation in nonprofits	Provide practical conceptualizations of the central objectives of evaluations and propose a framework that can guide negotiation processes	Netherlands	Review	It presents the relationships between the evaluation purpose, evaluation question, and the different levels of effects that should be measured. The selection of the evaluation method is contingent on the choices made within this framework.
Lynch-Cerullo, 2011	Moving from Outputs to Outcomes: A Review of the Evolution of Performance Measurement in the Human Service Nonprofit Sector	To examine the field-level pressures facing human service organizations and review the research on non-profit level responses to these pressures	USA	Review	After an examination of key innovations in social measurement, including the theory of change logic model, outcome standardization projects, and trends in calculating social value, as well as lessons learned from data-driven social innovation efforts, future directions in research and practice are proposed.
Macaulay, 2017	Conceptualizing the health and well-being impacts of social enterprise: a UK-based study	Examine how social enterprises portray their impact and how such impacts may be considered in health and well-being terms	UK	Primary	Revisiting previous conceptualizations in the extant literature, this work presents an 'empirically-informed' conceptual model of the health and well-being impacts of social enterprise-led activity. It thus presents a significant advance on previous hypothetical, theoretically based conceptualizations. It is considered that these findings further improve our overall knowledge of ways in which social enterprise and other parts of the

					third sector could be considered as potentially valuable 'non-obvious' public health actors.
Mackinnon & Stephens, 2010	Is Participation having an impact? Measuring progress in Winnipeg's inner city through the voices of community-based program participants	Develop indicators to measure the results of community-based organizations' work; describe key themes emerging from the research and the implications for social policy; reflect on PAR framework	Canada	Primary	Development of indicators that better reflect outcomes for participants
Manetti, 2014	The Role of Blended Value Accounting in the Evaluation of Socio-Economic Impact of Social Enterprises	to explore the role of blended value accounting (BVA) and SROI on theories of accounting, and especially within the context of measuring socio-economic impact of social enterprises	Italy	Review	BVA and SROI play a role with respect to eh positivist, critical, and interpretative theories of accounting
Mannarini et al, 2018	A Psychosocial Measure of Social Added Value in Non-profit and Voluntary Organizations: Findings from a	To consider a measure of social added value (SAV) and examine the validity of the measure presented in previous research	Italy	Primary	Shared member values were positively associated with SAV and that there were differences among organizations according to their legal forms and the organizational roles available

	Study in the South of Italy				
Medina-Borja & Triantis, 2007	A conceptual Framework to Evaluate Performance of Non-Profit Social Service Organizations	To provide a conceptual framework that can be used for the design and implementation of an integrated performance measurement system for non-profit organizations.	USA	Commentary	Using a conceptual framework in conjunction with data envelopment analysis offer a well-rounded evaluation of performance measurement
Medina-Borja & Triantis, 2011	Modeling Social Services Performance: a four-stage DEA approach to evaluate fundraising efficiency, capacity building, service quality, and effectiveness in the non-profit sector	To describe Data Envelopment Analysis (DEA) as a performance measurement tool/framework for multi-chapter non-profit social service organizations. Aspects of framework include fundraising efficiency, capacity building, service quality, effectiveness (outcome achievement). Also includes a case study	USA	Primary	Social service non-profits have a more challenging time being efficient in fundraising than under any other tenet. Furthermore, efficient fundraising is not a guarantee for efficient and high-quality service delivery, or client outcome achievement (effectiveness)
Millar, 2013	Social Return on Investment (SROI) and Performance Measurement	analyze survey and interview data and present how SROI is used and understood in	UK	Primary	Despite being accepted as an internationally recognized measurement tool for social enterprise, SROI is underused and

		health and social care settings			undervalued due to practical and ideological barriers
Minich, 2006	Can Community Change be Measured for an Outcomes-Based Initiative? A Comparative Case Study of the Success by 6 Initiative	To describe and compare the methods of evaluating community outcomes between chapters of the United Way's Success by Six initiative, and to describe challenges faced in the evaluation and communicating impact.	USA	Primary	Not all Success by Six chapters evaluated community outcomes, but similar measurement strategies are reported. Employee dissatisfaction with the logic model as a framework, though no preferred alternate method is identified.
Mio & Fasan, 2015	The Impact of Independent Directors on Organizational Effectiveness in Monetary and In-Kind Stakeholder Dialogue Museums	To test the existence of two different stakeholder dialogue's (monetary and in-kind) and test the impact of the number of independent directors on fundraising activity	Italy	Primary	NR
Mitchell, 2016	Evaluation and evaluative rigor in the nonprofit sector	Identifies catalysts and obstacles to evaluation activity and the correlates of evaluative rigor among US non-profits	USA	Primary	Results reveal that the most important catalysts to evaluation include the desire to improve program effectiveness and legitimacy, while the most important obstacles include insufficient time and money. Moreover, regression analysis finds that evaluation appears

					to be most rigorous when (1) evaluation is a priority, (2) a supportive organizational culture exists, (3) management requires evaluation, (4) evaluation is not primarily motivated by personal interest, and (5) evaluation is likely to reveal success. Overall, intrinsically motivated evaluation appears to be more rigorous than externally mandated evaluation, suggesting that stakeholders should work to help capacitate receptive non-profits to improve evaluative rigor instead of imposing external requirements.
Mitchell, 2018	Evaluation of nonprofit organizations: an empirical analysis	Proposes a set of hypotheses concerning the interrelationships between organizational characteristics and various aspects of non-profit evaluation	USA	Primary	Analysis reveals that although higher levels of staff compensation support many aspects of evaluation, higher levels of executive compensation exert negative effects. Additionally, evaluation culture mediates the effects of several variables on evaluation rigor and frequency. Practical implications are discussed for scholars and practitioners.
Molecke & Pinske, 2020	Justifying Social Impact as a form of impression management: Legitimacy Judgments of	To discuss forms of judgment by which we see social impact as legitimate. 2 main ones: cognitive and evaluative forms of judgment	Global	Primary	evaluative judgment is a better measure of legitimacy (how well the organization meets the audience goals as compared to multiple organizations) than cognitive judgment (the role of the organization

	Social Enterprises' Impact Accounts				matches the schemas in your head about similar actors' roles
Molecke, 2017	Accountability for social impact: A bricolage perspective on impact measurement in social enterprises	To investigate how social entrepreneurs handle the increasing pressure to measure social impact with formal methodologies through a bricolage lens	Global	Primary	These findings show how social enterprises combine material and ideational bricolage as well as seek to delegitimize formal methodologies to increase the legitimacy of their bricolage approaches for social impact measurement.
Mook, 2015	Measuring Social Enterprise Value Creation: The Case of Furniture Bank	To present a case study that explores the creation of value by a social enterprise, Furniture Bank, for its stakeholders	USA	Primary	This article calls for an integrated approach to social return on investment processes, incorporating both conventional accounting and social accounting.
Morales, 2018	Performance Measurement in Tribal Home Visiting: Challenges and Opportunities	To examine performance measurement challenges and opportunities in the context of tribal communities, and provide recommendations that may inform policy on performance measurement design	USA	Primary/Review	This article contributes to the literature by examining performance measurement challenges and opportunities in the context of tribal communities and provides recommendations that may inform future policy on performance measurement design and implementation in tribal communities.

		and implementation in tribal communities			
Moxham, 2013	Measuring up: examining the potential for voluntary sector performance measurement to improve public service delivery	To examine and challenge the notion that the process of measuring the performance of voluntary sector public service providers has the potential to improve the quality of public services	UK	Primary	The findings suggest that performance measurement is primarily being used to demonstrate compliance and not as a tool for service improvement.
Ormiston, 2019	Blending practice worlds: impact assessment as a transdisciplinary practice	Draws on practice theory to understand the purposes of impact assessment and how it influences and is enacted in, everyday organizational activities	UK; Australia	Primary	Impact assessment should be understood as a transdisciplinary practice evolving from multiple practice worlds such as strategy, accounting, marketing, and organizational learning
Pasupathy & Medina-Borja, 2008	Integrating Excel, Access, and Visual Basic to Deploy Performance Measurement and Evaluation at the American Red Cross	Used American Red Cross (ARC) as an example of using Data Envelopment Analysis (DEA) for performance measurement in a national non-profit organization. Also used Microsoft Excel and other basic applications to complete analysis.	USA	Primary	Designed and developed a system that uses DEA to make resource allocation recommendations to help Red Cross managers evaluate performance of various national chapters. Also discuss the useability of Microsoft Excel for this kind of performance measurement
Patton, 2001	The CEFP as a model for	Identifies and discusses the	USA	Primary	NR

	integrating evaluation within organizations	innovative contributions of the Collaborative Evaluation Fellows Project (CEFP) as a model for evaluation training and organizational development			
Pennerstorfer, 2019	Measuring Growth of the Nonprofit Sector: The Choice of Indicator Matters	Gives an overview of commonly used growth measures in existing non-profit literature, discusses informative value of various measures	Austria and Scotland	Primary/Review	Correlations between measures like number of organizations, income/expenditures, and assets are small. Researchers measuring the growth of the non-profit sector should be clear about the properties of their selected measure, and where possible should present alternate measures in their analysis
Polonsky, 2016	Perspectives on social impact measurement and non-profit organizations	To examine how Australian and US managers of NPOs and foundations view the measurement of the social impact of NPOs.	Australia and USA	Primary	Objective measures of impact are desirable, but recognized the difficulties in developing objective assessment frameworks enabling comparisons across the non-profit sector; there is an opportunity to reposition reporting expectations -- the NPO sector can pool together and build on each other's strengths and market their outcomes as a collective entity
Reid and Karambaya, 2009	Impact of Dual Executive Leadership Dynamics in	To examine several arts organizations in Canada and their	Canada	Case Study	NR

	Creative Organizations	leadership types and its impact on fundraising.			
Savaya & Waysman, 2005	The Logic Model: a tool for incorporating theory in development and evaluation of programs	Presents potential uses of the logic model tool in explicating program theory for a variety of purposes throughout the life span of programs; assessing the feasibility of proposed programs and their readiness for evaluation, program development, developing performance monitoring systems, building knowledge	Israel	Primary/Review	NR
Schatterman & Bingle, 2017	Government funding of arts organizations: impact and implications	Determine to what extent arts organizations are dependent on government grants and implications of government funding on other revenue such as individual and foundation support	USA	Primary	Arts organizations that receive state-level public funding are more successful in gaining donations from other sources
Schuh 2006	A framework to assess the development and capacity of	Present a framework to assess non-profit agencies' levels of	USA	Review	A framework to assess non-profit agencies' levels of development and capacity is presented. The framework is useful for evaluation because it

	nonprofit agencies	development and capacity			helps to predict and explain program implementation. The framework is useful for program planning because planners must select organizations that are suitable to implement programs, and sometimes they must build organizations' capacity to do the work. A scoring rubric called the stages of organizational change measures development and capacity, using the maturity modeling approach. Maturity modeling is a technique that emerged from the worlds of business and technology. 'Maturity' does not imply value judgments about organizations; rather, the level of development and capacity should fit the services to be delivered. The scoring rubric has good reliability and validity and has been applied in three ways: (1) to assess agency capacity at a single time point; (2) to describe capacity development; and (3) to evaluate capacity building initiatives.
Sirgy & Cornwell, 2001	Further Validation of the Sirgy et al.'s Measure of Community Quality of Life	To further test a measure of community quality of life (QOL)	USA	Primary	Global community satisfaction, in combination with satisfaction with other overall life domains (work, family, etc.), affects global life satisfaction.
Sirgy, 2000	A Method for Assessing	To develop and test a method for assessing	USA	Primary	The results provided support to the model and the assessment method

	Residents' Satisfaction with Community-Based Services: A Quality-of-life perspective	residents' satisfaction with community-based services			and measures used. From a managerial perspective, we showed how the model and the assessment method can be used by community leaders to tap citizens' perception of community quality-of-life and its determinants, identify strategic gaps or problem areas, and take corrective action.
Szijarto et al., 2018	On the evaluation of social innovations and social enterprises: Recognizing and integrating two solitudes in the empirical knowledge base	Presents findings from a systematic review and integration of 41 empirical studies on evaluation in social innovation (SI) contexts.	NR	Review	Differences of interest in collaborative approaches and facilitation of evaluation use
Taylor & Taylor, 2013	Performance measurement in the Third Sector: the development of a stakeholder-focussed research agenda	Develop the conceptual foundation of performance measurement (PM) in the Third Sector; derive a research agenda that provides a platform for future work	UK	Review	No extant models or frameworks are particularly suitable in this context, due in part to the wide range of stakeholders and the distinctive characteristics of Third Sector organizations (TSOs)

Turbide & Laurin, 2009	Performance Measurement in the Arts Sector: The Case of the Performing Arts	Address the issues of performance measurement for art not-for-profits	Canada	Primary	Performance measurements place as much emphasis on non-financial indicators as they do financial.
Unceta et al., 2016	Social innovation indicators	Provides the results of the Regional Innovation Index (RESINDEX), a conceptual and empirical model that explores indicators of social innovation at the organizational and regional level	Spain	Primary	NR
Veulli, 2018	Performance measurement in non-profit theatre organizations: the case of Greek municipal and regional theatres	Examine the implementation of performance measurement systems in Greek theatres using a qualitative case study strategy	Greece	Primary	NR
White, 2018	A Cook's tour: towards a framework for measuring the social impact of	Focus on organizations whose remit is to alleviate social problems, and therefore have a social purpose; contributes to the	NR	Primary	NR

	social purpose organizations	research gap by developing a framework for measuring the impact of social purpose organizations			
Wing, 2004	Assessing the Effectiveness of Capacity-Building Initiatives: Seven Issues for the Field	To address/engage with common issues in the evaluation of capacity-building in philanthropic organizations	USA	Review	Seven issues with capacity-building evaluation are identified and discuss with regard for implications for design, management, and evaluation of capacity-building initiatives

Appendix C: Grey Literature Resources

Organization	Country	Contact	Type of Organization (art, culture, leisure, technological, health care)	Relevant Resources or Documents
Call Hub	USA	NR	Communication company	Provides an in-depth outline of effective community outreach strategies. Provides a step-by-step description of how to develop a strategy to engage with a community.
Canada Council for the Arts	Canada	1-800-263-5588 OR 615-566-4414 info@canadacouncil.ca	Arts organization, funds art projects and programs.	The URL provided links to a page outlining the Council's assessment criteria for organizations or individuals applying for grants. This page discusses the application and decision-making process, which could be useful for the DVSA in deciding what they might want to measure, or which

				results they may want to demonstrate.
Canadian Housing and Renewal Association	Canada	613-594-3007 info@chra-achru.ca	Housing - national non-profit organization dedicated to supporting and strengthening the social housing sector; mission is to ensure all Canadians have an affordable, secure and decent place to call home	Paper - This paper draws from a workshop held at the Canadian Housing and Renewal Association's (CHRA) 2015 Congress on Housing and Homelessness. The workshop, Social Impact Measurement: Making the Case for Housing, focused on the how and why of measuring the social impact of investments in affordable housing from three different perspectives: a housing provider, a provincial government agency and a financial institution.
Charity Village	Canada	1-800-610-8134 help@charityvillage.com	Advisor - Career resource to over 170,000 charitable and nonprofit organizations across Canada. Canadian nonprofit sector's	An article that describes how NPOs can benefit from measuring their impact, the theory of change, and how to present the impact

			largest and most popular online resource for recruiting, news and how-to information.	that you have measured.
Chartered Professional Accountants (CPA) Canada	Canada	416-977-0748 member.services@cpacanada.ca	Finance - national organization established to support unification of the Canadian accounting profession	<p>“Performance Measurement for Non-Profit Organizations”</p> <p>This guidance document, aimed at senior management of NPOs, outlines one possible approach to strategy execution—a four-step framework called performance measurement for non-profits (PM4NPO) which utilizes a balanced scorecard methodology developed by Dr. Kaplan and Dr. Norton from Harvard University.</p>
Harvard Business School	USA	617-627-5104 alnoor.ebrahim@tufts.edu (Author of paper)	University/Education	Paper that describes a contingency framework for measuring social

				performance, suggesting that some organizations should measure long-term impacts, while others should focus on shorter-term outputs and outcomes. Also references the theory of change.
Jenny Larsson and Joan Kinnunen	Sweden		University/Education - Thesis	Thesis - Overviewing performance measurement in non-profits and whether it is necessary, as well as how to do it.
McKinsey & Company	USA; Europe	https://www.mckinsey.com/industries/social-sector/contact-us	Business advisors - Advises on strategic management to corporations, governments, and other organizations Provides strategy and management consulting services, such as providing advice on an acquisition, developing a plan to restructure a sales force, creating a new business strategy or	Article discussing measuring performance in NPOs. Describes a framework, “family of measures”, as well as measuring the success of an organization’s mission.

			providing advice on downsizing	
Monitoring and Evaluation in Art for Social Change	Canada		An arts and social science based national project with international partners. The project was made possible by a Social Sciences and Humanities Research Council of Canada (SSHRC) Partnership Grant.	A list of existing evaluation toolkits and guides supplied here: Evaluative research mini course
Mowat Research	Canada	416-946-8900 munkschool@utoronto.ca	University/Education	Paper - Strategy to offer the platform for charitable and nonprofit organizations to collaborate and align priorities on outcomes measurement in Canada. This paper explores current barriers to measuring outcomes and what is needed to create an enabling environment for Canada's charitable and nonprofit sector to participate more readily in outcomes

				measurement activities.
National Endowment of the Arts	USA	202-682-5400 webmgr@arts.gov	A national organization that offers updates on arts of all media, and bridges the gap between research and arts as recreation in the U.S.A.	Many resources on evaluation found here Or you can perform the Search function on the website and search “evaluation” for a wider variety of articles.
Of/By/For/All	USA	admin@ofbyforall.org	A national organization that focuses on making organizations of, by, and for diverse communities -- incorporating the voices of the community into how the services are provided.	Includes resources from the organization, including a guide on respectful audience surveying, also available here
Ontario Arts Council	Canada	416-961-1660 OR 1-800-387-0058 (toll-free in Ontario) info@arts.on.ca	An agency primarily focused on arts, operates “at arm’s length” from the Ministry of Heritage, Sport, Tourism and Culture Industries. Directed by 12 volunteer board members from communities throughout the	Good resource that includes grants that the DVSA may be eligible for

			province and invests in Ontario-based artists and arts organizations.	
Sopact	USA	1-510-676-9502 https://www.sopact.com/company/contact-us	A consulting agency that works with social purpose agencies to develop evaluation plans and demonstrate their social impact.	Includes downloadable resources , and digestible videos discussing benefits and challenges of social impact measurement
The Canadian Network for Arts & Learning	Canada	info@eduardarts.ca	Arts organization, but specifically a national organization that works to demonstrate the value of the arts and learning. Has a network model for stakeholders like the DVSA, and acts as a political interest group to influence policy surrounding art and learning.	A handbook they commissioned on assessing the impacts of arts and learning: A map or arts and learning organizations across the country, with opportunity to contribute to the map
Thesis from Universitat Jaume I	Spain	https://www.linkedin.com/in/f%C3%A1tima-bartoll-andreu-95516311b/?originalSubdomain=es	University/Education	Thesis paper - The purpose of this paper is to deepen the study of the measurement of social impact, its strengths and weaknesses.

VolunteerHub	USA	1-877-482-3340 info@volunteerhub.com	Volunteering	Article - Discussing nonprofit metrics to measure during impact measurement.
W.K. Kellogg Foundation	USA	269-968-1611 https://www.wkkf.org/contact-us	A U.S. national nonprofit that prioritizes services for vulnerable children, as a way to building stronger families and equitable communities.	Includes a handbook on evaluation. includes a logic model development guide (which can underpin a theory of change)
Yukon School of Visual Arts (report of evaluation)	Canada	NR	This is a report on the evaluation of a Bachelors program in the fine arts at the Yukon school of Visual Arts. The report includes evaluation methods and a framework. Some of the long-term outcomes/impacts include contributions to cultural industries, including First Nation visual art, in Yukon and in Canada.	See URL for report.