1-2019

Performance Measurement in Maine’s Social Sector: The Intersection of Philanthropy and Nonprofits

Susy Hawes
The Data Innovation Project

Sarah Goan
The Data Innovation Project

Performance Measurement in Maine's Social Sector: The Intersection of Philanthropy and Nonprofits

An inquiry into the evaluation and performance measurement expectations among funders and grantees

Data Innovation Project
January 2019

Foreword

The Data Innovation Project (DIP) is housed within the Cutler Institute of Health and Social Policy at the Muskie School of Public Service, University of Southern Maine. Founded in 2016, the DIP partners with organizations to increase their capacity to utilize data to clearly define their work, achieve results, and demonstrate impact in Maine. The DIP believes that all mission-driven organizations in Maine have the potential to use data to affect change to improve conditions for the people and places with and for whom they work. Over the past two years, the DIP has engaged with over 250 Maine nonprofits who are navigating the varied data reporting requirements of a diverse portfolio of funders with varying degrees of success and challenges.

With the support of the Maine Philanthropy Center, the DIP undertook this project to better understand foundations' expectations of grantees in terms of performance measures and outcomes as outlined in their grant applications. The DIP is pleased to share this report to begin a conversation within Maine to look more closely at how this important relationship between foundations and nonprofits may benefit from increased understanding and clarity around measurement approaches so that both dollars and efforts lead to meaningful and measurable results across the state, and ultimately enhance the health and well-being of all Mainers.

Acknowledgements

The Data Innovation Project is grateful for the support and collaboration of Barbara Edmond and the Maine Philanthropy Center (MPC). MPC provided crucial support and collaboration on this project by making their 2017 *Directory of Maine Grantmakers* available to the DIP and answering questions about the database.

We are additionally thankful for the hard work of our graduate assistants, Anna Korsen and Hannah Whitney, who put many hours into digging for, sorting through and analyzing the applications to make this report possible. This project was spurred by our conversations with many in Maine's nonprofit and philanthropic sector, whose honesty, experience and commitment will continue to move this work forward.

UNIVERSITY OF SOUTHERN MAINE

CUTLER INSTITUTE
MUSKIE SCHOOL OF PUBLIC SERVICE

DATA INNOVATION PROJECT

Table of Contents

Introduction and Purpose

Maine's philanthropic and nonprofit sectors are woven into the fabric of the state's prosperity. In 2017, there were 6,875 registered 501(c)3 nonprofits in the state of Maine who contributed $12.6 billion to the economy through wages paid, retail and wholesale purchases, and professional services contracts.[1] Nonprofits employ 1 in 6 workers, which is larger than any other industry in the state.[2]

Maine's foundations are a tremendous asset to the state and essential to the success of nonprofits. In 2016, there were 356 foundations registered in the state that granted $175 million throughout the world and held assets of $2.9 billion.[3] Yet, national research has found that both foundations and nonprofits struggle to meaningfully integrate data and evaluation into that story of success; nationally, 94% of nonprofit professionals do not feel they are using their data effectively and 75% of foundation professionals do not feel evaluations provide meaningful insights.[4] Performance measurement and evaluation practices between and within foundations and nonprofits is a topic of growing interest that has been consistently raised both locally and in national research and publications.[5] There is an emergent focus on ensuring that nonprofit grantees are accountable to their impact on the wellbeing of people and communities.

Awareness of what data foundations are requesting and how those data are used is integral to understanding the impact of Maine's nonprofits. This paper explores this important topic with the goal of providing the philanthropic and nonprofit sectors with Maine-specific information intended to guide data-informed conversation and decision-making. The primary question that drives this inquiry is:

> ***How do foundations who fund in Maine expect their grantees to measure change, in terms of performance measures and outcomes, as outlined in their grant applications?***

The intent of this project was to gather **baseline** data about what foundations who fund in Maine are asking of nonprofit grantees by reviewing foundations' funding applications. While this application scan does not represent all foundations that fund in Maine and the findings cannot be generalized, our hope is

---

[1] Maine Association of Nonprofits. (2019 January). *Adding Up Impact: Maine Nonprofits at Work*. Retrieved from http://www.nonprofitmaine.org/about-nonprofits/nonprofit-research/adding-up-impact-report-2019/

[2] Ibid.

[3] Maine Philanthropy Center. (2018). *2018 Giving in Maine Report*. Retrieved from https://www.mainephilanthropy.org/2018-giving-reports

[4] Everydayaction. (2016). *The State of Nonprofit Data: The Good, the Bad, & the Ugly.* Retrieved from https://act.everyaction.com/2016-nonprofit-data-whitepaper

[5] Buteau, E., Glickman, J., Loh, C., Coffman, J., & Beer, T. (2016 September). *Benchmarking Foundation Evaluation Practices. Retrieved from* http://research.cep.org/benchmarking-foundation-evaluation-practices; Evans, R., Kasper, G., & Siesfeld, T. (2017, February 1). *Moving Toward a Better Future for Measurement*. Retrieved from https://ssir.org/articles/entry/moving_toward_a_better_future_for_measurement; Gugerty, M.K. & Karlan, D. (2018 Summer). *Ten Reasons Not to Measure Impact—and What to Do Instead*. Retrieved from https://ssir.org/articles/entry/ten_reasons_not_to_measure_impact_and_what_to_do_instead?utm_source=Enews&utm_medium=Email&utm_campaign=SSIR_Now&utm_content=Title; Stanford Social Innovation Review for the Aspen Institute Program on Philanthropy and Social Innovation. (2012, May 17). *Advancing Evaluation Practices in Philanthropy*. Retrieved from https://ssir.org/supplement/advancing_evaluation_practices_in_philanthropy; Tatian, P. (2016, March 15). *Performance Measurement to Evaluation. Retrieved from https://www.urban.org/research/publication/performance-measurement-evaluation-0*

that this report will spur thoughtful, data-informed conversations both between and within foundations and nonprofits regarding **outcome** measurement and reporting.

The remainder of this report is organized into the following sections: methodology (sample and analysis), findings, and recommendations and best practice considerations. Appendices provide further details on methods, recommendations for further reading and research, and a glossary of terms.

*Definitions for words that are **bolded in purple font** can be found in Appendix C.*

Methodology

The primary data collection effort in this inquiry involved collecting and reviewing **evaluation**-related questions from grant applications from a sample of foundations who have a history of giving in Maine. This project did not examine the post-application period to see how foundations may address performance measurement or evaluation activities once grants are awarded. The DIP reviewed the Maine Philanthropy Center's 2017 *Directory of Maine Grantmakers* that includes profiles of almost 400 grantmaking foundations based in Maine, identify Maine as a funding location, and/or have a history of funding nonprofits in Maine. The directory contains a breadth of information related to each of the foundations, spanning from the contact information of the grants administrator, to information about the foundation's giving foci, the assets and number of grants paid by the foundation, and samples of grants given in Maine by the foundations, among other indicators. The directory was pivotal to narrowing the sampling approach and obtaining information to supplement findings.

Sample

The following criteria were used in the sampling approach for this investigation:

1. ***Foundations that are currently accepting applications***. At the time of review, of the almost 400 grantmaking foundations included in the directory, 199 were currently accepting applications (approximately 50%).
2. ***Foundations that have a website, potentially where the application for funding could be found.*** Since the majority of foundations have no or limited staff, this criterion was chosen to minimize the burden of contacting foundations. Of the 199 foundations that were currently accepting applications in the database, 102 had a website included in the directory (approximately 25% of the total grantmaking foundations).
3. ***Foundations that have large amounts of money paid in grants or large number of grants given.*** This criterion was selected to prioritize data collection from those foundations that have a large financial presence in terms of giving throughout the state.

Applications were collected from November 2017 through March 2018. In total, 75 applications were accessed and met the criteria to be included in the analysis. In some cases, one foundation had multiple grants with distinct applications. When more than one application within a foundation asked unique evaluation-related questions, individual applications were included in the analysis. The 75 applications included in the analysis for this project come from 65 foundations. The 65 included a mix of funders, including, for example, corporate, family, private and public funders. Total foundation assets of those included in the analysis ranged from less than $100,000 to more than $100 million. Total amount of money paid in grants ranged from less than $30,000 to more than $15 million. The total number of grants given ranged from less than 10 to more than 1,000.[6]

Snapshot of foundations in sample

Average total assets	**$35,780,555**
Average total $ paid	**$2,535,932**
Average total # of grants given	**151**

Analysis

Each of the 75 applications were assessed for eight content areas, detailed below. These areas were selected because they represent key elements of evaluation and performance measurement and impact the grantees ability to appropriately plan for, budget, and follow-through on tasks. In addition, these content areas help gauge the extent to which foundations are concerned with and support grantees' performance measurement and evaluation processes.

- Logic Model or Theory of Change
- Outputs
- Defined, measurable outcomes
- Evaluation plan
- Foundation outcomes
- Evaluation support
- Progress report
- Associated Grant Makers Common Proposal

Content areas were either rated as yes or no (present or not present) or on a scale of one to four where one indicated the area was not present, two that it was emerging, three that it was "on the way" and four that the area was fully present in the application. More details on ratings, including examples of scores, are discussed in the findings section. The applications were reviewed systematically to identify the extent to which the area were present. To ensure inter-rater reliability, two researchers independently reviewed 10 applications and then compared the results. This process identified instances in which the ratings were applied inconsistently. This reliability check was completed once more. Once reconciled, one researcher continued rating the applications independently. For a more detailed explanation of how these values were assigned and calculated, please see Appendix A.

In addition to scoring content areas individually for each application, a total or comprehensive score was computed to provide a higher-level view of the overall emphasis on performance measurement or evaluation in the reviewed applications. Scores were computed by combining the individual scores for **outputs**, outcomes and evaluation plans.

---

[6] These are estimated minimums and maximums intended to maintain the anonymity of foundations included in this analysis. Any outliers have been removed.

Performance Measurement vs. Evaluation: What's the difference?

It is important to understand the difference between *performance measurement* and *evaluation* as they are two very distinct ways of looking at the work of organizations. While both performance measurement and evaluation contribute to the evidence of program effectiveness and may be used to promote organizational learning about what works, there is a fundamental difference in what each is trying to accomplish.

Performance measurement is a responsive and adaptive approach to making sure a program is on track to meet its goals using inputs, outputs and outcomes that are attributed to the program. It is typically ongoing and routine and can be conducted by program staff to inform service or program delivery.

Evaluation is rooted in research design methodologies and is set up to address a set of pre-determined questions (or hypotheses) to test some aspect of whether the program's theory (its expected outcomes) held true. Evaluation considers the question *would you see the same results without the program?* This implies conducting more rigorous research procedures, such as a randomized control trial (RCT). Evaluation generally occurs less frequently and is conducted with the guidance or support of an independent professional.

Tatian, P. (2016, March 15). *Performance Measurement to Evaluation. Retrieved from* https://www.urban.org/research/publication/performance-measurement-evaluation-0

Content Areas: Guiding Questions and Ratings

Logic Model or Theory of Change: Organizational logic models, which may include a theory of change, define organizations' strategies and outputs and match these to short, mid and long term outcomes. Organizations that have an organizational and/or programmatic logic model are more likely to have clarity around expected outcomes to report to funders. Similar models are called a theory of action or results action plan.

Does the foundation ask grantees to share a logic model and/or theory of change?

Rated as either yes or no.

Outputs: Outputs are numbers and counts used to show an immediate measure of what the program did. They help programs to ascertain the extent to which they have completed the work they had planned (e.g. # of people served, # of students enrolled, # of newsletters distributed).

To what extent does the foundation ask grantees to define their program outputs?

Rated 1-4.

Defined, measurable outcomes: Outcomes identify the changes that are expected as a result of the implemented strategies or activities and can be separated into the short-, mid- and long-term influence of a program or initiative.

To what extent does the foundation ask grantees to define their short-, mid- and/or long-term outcomes?

Rated 1-4.

Evaluation plan: Evaluation plans describe the methods and tools that will be used to measure intended outcomes. The term "evaluation plan" is intended to capture both performance measurement and evaluation planning.

To what extent does the foundation ask grantees to define their evaluation plan?

Rated 1-4.

Foundation outcomes: Foundation outcomes are defined for this project as larger, measurable, population-level outcomes. Foundations may ask grantees to identify the outcome to which their program or initiative is contributing. For the purposes of this project, these are different from a foundation's overarching giving focus in that they are defined and measured at the population level by the foundation.

Does the foundation ask grantees to identify the foundation outcome(s) to which their work contributes?

Rated as either yes or no.

Evaluation support: Foundations may indicate in an application if they provide grantees with, or direct them to, resources or supports specifically related to evaluation or performance measurement.

Does the foundation provide grantees with evaluation resources or supports?

Rated as either yes or no.

Progress report: Foundations may indicate in an application or in their application guidelines whether or not they require a progress report be completed during or at the end of the funding cycle. Progress reports that ask about progress toward outcomes keep grantees accountable to the outcomes they identified in their original application.

Does the application indicate a progress report is required?

Rated as either yes or no.

Associated Grant Makers Common Proposal: The Associated Grant Makers Common Proposal Form is used or accepted by some foundations. The AGM form was developed "to make the grant application process easier for both grantmakers and grantseekers" and includes basic questions on evaluation.[7]

Does the foundation use the AGM common proposal form?

Rated as either yes or no.

[7] Associated Grant Makers. AGM Common Proposal Form. Retrieved from http://www.agmconnect.org/resources-tools/resource-center-agm/agm-common-proposal-form.

Findings

While the extent to which foundations asked grantees evaluation and performance measurement questions varied between individual foundations, in general this analysis indicates there is minimal direction or clarity from foundations for grantees to consider the various aspects of performance measurement or evaluation as they undertake a funded project. When present, the questions regarding measurement were often vague, lacking details around expectations, and used language and terms inconsistently.

Only four of the 75 applications reviewed accepted some type of common grant application and there was extreme variety in the ways and extent to which questions were asked. This amount of variability may limit the capacity of nonprofits in Maine to undertake a consistent or intentional measurement approach for programs or initiatives that have multiple funders (or when they are applying for grants from various funding organizations). In addition, lack of clarity in application questions suggests there may be misalignment between what nonprofits understand is expected of them and what foundations are actually expecting. The absence of a common language across philanthropies and nonprofits may be inhibiting data-informed decision making across the sector. The following pages detail specific findings by content area.

Findings by Content Area

Logic Model or Theory of Change*: Overall, funders did not require grantees to present a theory of change or logic model in the grantmaking process.* Of 75 applications, three asked for a logic model or theory of change (4%). Of these three, one said including the logic model or theory of change was optional. Seventy-two applications (96%) did not ask for a theory of change or logic model.

Outputs: *The majority of funders did not ask for grantees to track outputs; those that did often only requested a count of the number of people served.* Specifically, 48 (64%) applications included in the sample did not ask grantees to identify any outputs they would track as a result of the grant funding. Of the remaining 27 applications, 23 (31%) were rated 2 or "emerging," three were rated 3 or "on the way" (4%), and one was rated 4 or "fully present." Of the 31% of foundations that scored a 2, the most common output requested was the number of people expected to be served by the grant funding. Most foundations did not ask programs to consider more specific programmatic measures that would help to determine the extent to which grantees will complete the work they had planned under the grant funding.

Defined, measurable outcomes: *While more than half asked for some sort of outcome-type measurement, funders used various terms to elicit outcome-type measures from grantees and few provided examples.* A third (33%) of applications did not ask any questions related to outcomes, while 30 applications (40%) scored a 2, indicating outcome-related questions were emerging in the application. Twelve (16%) scored a 3 and 8 (11%) scored a 4.

In asking about outcome-type measures, funders used a variety of terms. Some foundations used the term outcome; others used goals, objectives, **impact**, anticipated **results**, or expected accomplishments. Although the terms appeared to be used interchangeably across and at times within applications,[8] each had its own distinct definition in the context of evaluation or performance measurement. Very few applications provided any examples to help grantees write outcome statements. In one that did, the example was incorrect.[9]

Although in some applications it was clear that the outcomes should be measurable, the majority asked an overarching question about the anticipated outcomes without detailing stipulations or questions regarding if and how the outcomes should be measured. Fewer than five provided any time frame for grantees to outline anticipated outcomes; for example, will the outcomes be accomplished in the short term, mid-term or long term?

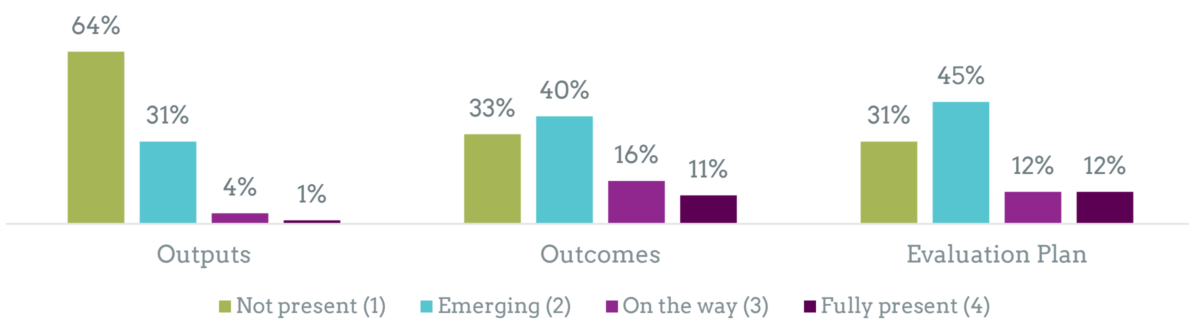
Evaluation Plan*: While more than two-thirds of funders asked some type of question about measurement or evaluation planning, the majority lacked clarity in the specific information they hoped grantees would provide.* Thirty-four applications (45%) scored a 2, indicating questions about an evaluation or performance measurement plan were emerging in the application. Twenty-three applications (31%) did not ask grantees any questions regarding how they would evaluate outcomes or the success of the funded project; these applications did not ask for an evaluation plan either. Nine (12%) scored a 3 and 9 (12%) scored a 4.

Questions varied widely between applications; some applications asked what specific tools would be used to measure the success of the program while others simply asked, "How will the program be evaluated?" Other language used to indicate an evaluation plan included: "What is the method for evaluating project success?", "How will the effectiveness of the project be evaluated?", and "What is the process for measuring success?" Other terms used to describe an evaluation plan included a statement of evaluation activities, methods to measure and/or **monitor** progress to results and measures of effectiveness. No applications clearly defined expectations related to evaluation versus performance measurement. There was a lack of clarity in the extent to which foundations were asking grantees to detail evaluation plans and the purpose for such questions.

[8] One application asked grantees to define their outcomes, but throughout the application used the terms objectives, goals, impact and outcomes interchangeably.
[9] The application question asked for outcomes, but the example was of a program output.

Although applications consistently did not ask for outputs, almost two thirds asked grantees for outcomes and evaluation plans.

Percentage of applications scoring 1-4



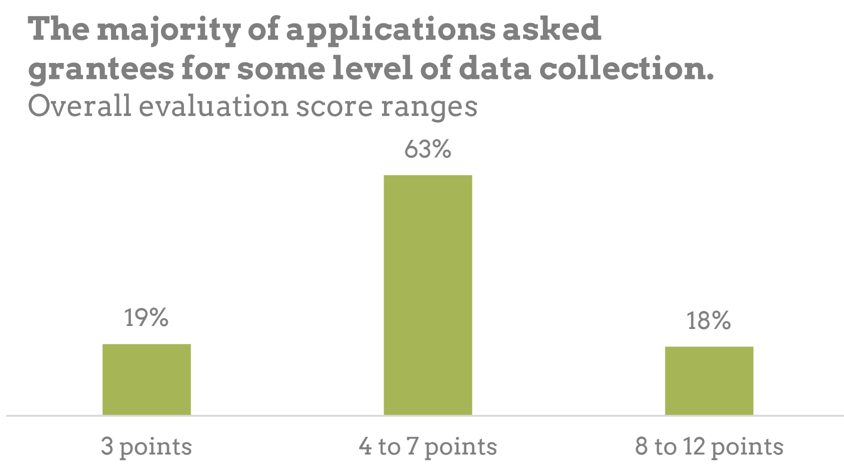
Foundation Outcomes*: The vast majority of funders did not ask grantees to identify how their efforts would fit under larger foundation-level outcomes.* Specifically, 68 applications (91%) did not ask applicants to demonstrate how they would help contribute to measurable foundation identified population-level outcomes. While most foundations asked questions about how the applicant would help to meet the larger goals of the foundation or the requirements of the grant, only seven (9%) applications asked grantees to identify the foundation outcomes to which they would contribute through the grant funding.

Evaluation Support: *In general, foundations did not indicate in applications that they offered any type of evaluation-related support to help build the data literacy or capacity of applicants.* Of the 52 applications that asked a question about an evaluation plan, three (6%) provided grantees some type of support regarding evaluation. Examples of support included an evaluation planning guide and a list of potential measurement tools.

Progress Report: *The majority of funders did not indicate that they required a progress report. In addition, there was some inconsistency in the extent to which foundations asked evaluation-related questions and what was expected in terms of progress reporting.* Forty-five applications (60%) did not indicate requiring a progress report; 30 (40%) did. Only four of the 14 (29%) applications that asked somewhat robust questions regarding outcomes and monitoring (as detailed in the overall evaluation score findings below) indicated they required a progress report. Conversely, of the 13 applications that asked no questions about evaluation, four (31%) required a progress report. In one instance, the grant application did not ask any questions related to an evaluation plan, but the progress report asked grantees to report on how the project was evaluated. Another application asked detailed questions about the defined, measurable outcomes the grantee anticipated achieving using the grant funds, but did not ask any questions about how they intended to evaluate progress toward such outcomes. These examples suggest a lack of alignment between what is detailed in the application and what is expected in progress reporting.

Associated Grant Makers Common Proposal: *Funders were not typically using common grant proposal forms.* Two (3%) of the reviewed applications accepted the Associated Grant Makers Common Proposal form. Two foundations, including one of those that accepts the AGM common proposal, accepted the Common Grant Application developed by Maine Philanthropy Center.[10] As previously mentioned, the AGM form contains basic evaluation questions.

Overall evaluation score: *The majority of funders asked for some level of data collection and tracking on outputs, outcomes or evaluation planning.* The overall score was compiled across all the areas and ranged from 3 points (lowest score) to 12 points (highest score). Of the 75 applications, 14 (19%) scored a 3, meaning they asked no questions about outputs, outcomes or evaluation or performance measurement planning. The majority of applications, 47 (63%), scored between a 4 and a 7, meaning they asked some questions about evaluation. Thirteen (17%) scored between an 8 and an 11, meaning they were more likely to ask more robust questions about outputs, outcomes and/or evaluation planning. Only one application was deemed to be at the highest level and scored a 12.



[10] Maine Philanthropy Center. (2016 September) Common Grant Application Information. Retrieved from https://www.mainephilanthropy.org/resources/common-grant-application-information.

Funding Evaluation Efforts

Although a detailed analysis of grant funding ranges was outside the scope of this research, grant ranges were collected when available in initial data collection efforts. Forty of the 75 applications (53%) included a dollar range or funding amount on the application or easily visible on the website. To determine if there is a correlation between the amount of funding available and the extent to which grantees are asked about evaluation, the amount of money available through the grant was compared to the overall evaluation score. It is important to note that because some grants were for specific dollar amounts (e.g. $5,000) and others were for ranges (e.g. up to $25,000) the specificity of this analysis was limited.

Of the applications that scored an 8 or above in their overall evaluation score, indicating they asked somewhat robust questions regarding outcomes and monitoring, eight of the thirteen included a grant range. One of the applications asked grantees to describe the specific measurement tools that would be used to "measure the success of the program," but were only funding grants up to $5,000. Three others with an overall score of 8 or above were for grants up to $25,000.

On the other end, applications for funding projects over $100,000 did not consistently ask for outcomes and/or evaluation plans. Of the seven applications reviewed that funded projects at or beyond $100,000, four had an overall evaluation score between 3 and 6, indicating limited demand in having grantees report on their ability to identify and track progress toward outcomes.

It is important for grantmakers to consider the effort and time it takes for detailed measurement planning and reporting. Although a thoughtful performance measurement plan with allocated funding attached to the work might be appropriate for a $25,000 grant, these activities would take up a large percentage of smaller grants (or need to be undertaken by leveraging other organizational resources).

More broadly, this inquiry found that the application processes vary widely. An area for further research would be to look into the time and resources it takes nonprofits to apply for funding in comparison to the amount of funding available.

Recommendations and Best Practice Considerations

Evaluation and monitoring should not be done retroactively; quality efforts require a thoughtful and systematic plan that is in place from the beginning of an initiative. There are a number of opportunities for Maine's philanthropic and nonprofit sectors to better support meaningful performance management and ensure better outcomes for the people and communities of Maine. In general, improved clarity around reporting and evaluation expectations at the outset of applications is important for the success of both funders and grantees. Without this clarity, it will be challenging for grantees to appropriately plan, fund, and account for staff time and resources to undertake meaningful evaluation efforts. Streamlined grant application processes across Maine's philanthropic sector will ensure nonprofits can use resources efficiently and continue to have the greatest impact.

It is with these considerations in mind that the following recommendations and best practice considerations are offered to funders and nonprofits as an opportunity to continue the critical thinking around this topic.

1. ***Be realistic when detailing expectations and proposals for evaluation or performance measurement; such processes should be closely tied to the total amount of funding available.***
   Measurement expectations should be aligned with the amount of grant funding available; larger grants may expect more in terms of measurement while smaller grants should expect less. Findings indicate that there is a mismatch at times between the amount of grant funding available and the performance measurement and evaluation expectations. Small grants will likely only be able to report on outputs and some short term outcomes, while larger grants that are multi-year may be able to report on short to mid-term outcomes. Reporting on any outcomes will require both preparing grantees with the requirements ahead of time so they can plan, as well as allotting funding to monitoring and evaluation when necessary.

   a. ***With a consideration to grant size, consistently allocate funds from grant budgets to support performance measurement or evaluation efforts.*** Asking for grantees to track, monitor, and/or measure the outcome of their work as related to a specific funding source costs money. When considering budgets for monitoring or evaluation efforts, it is necessary for both funders and grantees to understand what evaluation and performance measurement are, how they differ and when it is appropriate to use each of these approaches. Large scale evaluation plans that involve on site **quasi-experimental designs** (QEDs) and **randomized controlled trials** (RCTs) are typically 15-20% of a project budget.[11] Even smaller performance measurement efforts such as developing a data collection plan, launching a **survey** or conducting a **focus group** take significant amounts of staff time and resources. Depending on the capacity of the nonprofit and the

[11] Corporation for National and Community Service. (2016, March 13). Evaluation Budgeting Quick Guide. Retrieved from https://www.nationalservice.gov/sites/default/files/documents/SIF%20Evaluation%20Budgeting%20Quick%20Guide.pdf.

plan for measurement, both funders and nonprofits should consider a distinct budget line in grant applications to support performance measurement or evaluation.

2. ***All reporting requirements related to performance measurement or evaluation should be clarified in the beginning of the application process so grantees are able to adequately budget and prepare.***
   Whether foundations require a detailed mid- or end-of-project progress report or a simple report to close out the project, any and all expectations around data collection and reporting should be clearly detailed during the application phase. Such clarity gives grantees a greater understanding of the extent to which the foundation will expect to see measurable qualitative or quantitative data related to the proposed outcomes at the end of the grant.

3. ***Ensure language used in applications related to measurement efforts is as specific and accurate as possible to help align expectations for all parties.***
   When asking grantees to describe an approach to an evaluation or performance measurement plan, the correct use of terms in the application is extremely important. Funders may consider including a glossary or specific examples for each of their evaluation-related questions.[12] Otherwise, funders may inadvertently ask grantees for a plan to undertake a thorough and multi-year evaluation to measure the long-term outcomes of the funded project or program when all they actually hoped for is a basic report of outputs over time.

4. ***Consider aligning language, definitions and expectations across the sector that goes beyond outputs and number of people served.***
   Aligned language and expectations intended to track impact on the wellbeing of those served by the sector has the potential to improve both program and population-level outcomes. Given the data literacy needs of the sector, a strategy to build common capacity, through a toolkit or a convening, and movement toward shared language and approaches to performance measurement across Maine would be an important contribution to the landscape.

5. ***Focus (and limit) data collection requirements to what will be used by funders for accountability and grantees for continuous quality improvement.***
   Reviewing these results is an opportunity for foundations to carefully consider how much they need to know about the outcomes of their funded projects. Tracking the number of training hours offered or the success of a series of workshops in a methodical way that produces meaningful and quality results requires the commitment and dedication of both nonprofit staff and funding organizations. If nonprofits are being asked to track data beyond simple counts, they should be provided with the funding to do so and a clear and meaningful reason why. Consider in what ways the foundation will use the results. Ask the question: *How can data collection efforts be structured in a way that not only provides the foundations with helpful information around accountability and allows them to be more data-informed in their giving approach, but also builds the capacity of nonprofits to integrate data into their organizations in a sustainable manner?*

[12] Please see Appendix C for a Glossary of Terms referenced in this report.

6. ***Consider how program monitoring and evaluation contribute to the aligned contribution of work toward common goals.***
   When organizations have a solid performance measurement plan that outlines their mission, short, mid and long term outcomes and strategies for achieving those outcomes, they are poised to develop a data collection and progress monitoring system. Without a clear measurement plan and related data to inform program decision making, it is difficult if not impossible to understand the extent to which the organization is having an impact in the community.

Funding organizations across Maine have an incredible opportunity to support Maine nonprofits in contributing to larger systems change. Financial support for efforts to implement, increase or improve performance measurement and evaluation will build the internal capacity of nonprofits to systematically evaluate their activities and make data-informed decisions to strengthen the impact of their work. In addition, when nonprofits have a strong, realistic data and evaluation plans, they are more likely to be able to articulate to funders their existing data collection efforts rather than defaulting to the creation of a new and possibly unrealistic plan.

This report is shared with the hope that it sparks a data-informed conversation about performance measurement and evaluation within and between funding organizations and nonprofits in Maine. As researchers and evaluators committed to building the data capacity of mission-driven organizations, the Data Innovation Project sees this as an opportunity for funders and grantees to strengthen their relationships and achieve greater impact. Beyond using performance measurement and evaluation to track individual programs and initiatives, a thoughtful approach to this process can benefit both nonprofits and foundations. For nonprofits, clarity in expected outcomes and **metrics** will streamline data collection efforts and applications. For foundations, strategic use of data from grantees can inform internal strategic planning processes. And collectively, **when organizations are using data to inform internal decision making, they are accountable not only to themselves and funders, but also to the people and places across Maine whom they serve.**

Appendix A. Findings by Area: Explanation of Values

Outputs: Applications were scored a 2 if the funder asked grantees for the total number of people expected to be served by the grant funding. Applications scored a 3 if the foundation asked the grantees for the total number of people to be served as well as additional counts related to the funding request; for example, the total number of jobs expected to be created or trainings expected to be offered. A score of 4 was used to indicate the foundation requested more extensive data collection of anticipated program outputs.

Defined, measurable outcomes: A score of 2 signified the application contained a question regarding outcomes (using this term or another), but the question was either contained within another question or the intensity with which outcomes were required was low. For example, an application scoring 2 may ask in one question for a narrative for the project, the expected accomplishments and a description of how they will be obtained. Applications were rated 3 if they clarified the outcomes should be measurable, for example and asked a separate question about how the outcomes would be measured. An example of an application that scored a 4 ask the grantee to describe the project objectives, the strategies they will use to achieve the objectives and the larger desired outcomes of the project.

Evaluation plan: Applications that scored a 2 indicate the presence of some type of evaluation question. These questions were often broad and many included a request for an evaluation plan within another application question. For example, "Indicate the criteria for success and the process for measuring success." Applications that scored a 3 were more thorough in asking for specifics about a process for measuring or tracking outcome data. For example, "How will you regularly measure, document, and evaluate your results, (both qualitative and quantitative)?" Applications scoring a 4 asked multiple questions pertaining to an evaluation plan, often including specific questions about what tools the grantee would use to evaluate their program and what data and information they would collect and from whom, for example. In addition to one or more evaluation plan-related questions, many of the applications scoring a 4 asked questions regarding whether or not the people served would be included in the evaluation process.

Appendix B. Additional Reading and Resources

Listed below are a number of opportunities for additional reading and resources. A short description of the article is followed by the citation.

This report uses infographics to highlight the results of a survey of senior evaluation and program staff from foundations across the United States and Canada that attempted to capture the internal evaluation practices of foundations. The report includes a guide with discussion questions for foundations to consider regarding evaluation.

Buteau, E., Glickman, J., Loh, C., Coffman, J., & Beer, T. (2016 September) Benchmarking Foundation Evaluation Practices. Retreieved from http://research.cep.org/benchmarking-foundation-evaluation-practices

This short article circles around three "defining pillars of a better future" for monitoring, evaluation, and learning. It discusses the relationship between foundations and nonprofits and how to meaningfully integrate monitoring and evaluation practices into grant making and management.

Evans, R., Kasper, G., & Siesfeld, T. (2017, February 1). Moving Toward a Better Future for Measurement. Retrieved from https://ssir.org/articles/entry/moving_toward_a_better_future_for_measurement

This article in the summer 2018 edition of the Stanford Social Innovation Review discusses and benefits and cautions of attempting an impact evaluation and uses four categories to detail when not to measure impact, and alternative strategies of performance measurement.

Gugerty, M.K. & Karlan, D. (2018 Summer) Ten Reasons Not to Measure Impact—and What to Do Instead. Retrieved from https://ssir.org/articles/entry/ten_reasons_not_to_measure_impact_and_what_to_do_instead?utm_source=Enews&utm_medium=Email&utm_campaign=SSIR_Now&utm_content=Title

This six article supplement to the Stanford Social Innovation Review discusses considerations, practices and principles for foundations who are undertaking efforts in program monitoring and evaluation, both internally and with their grantees.

Stanford Social Innovation Review for the Aspen Institute Program on Philanthropy and Social Innovation. (2012, May 17). Advancing Evaluation Practices in Philanthropy. Retrieved from https://ssir.org/supplement/advancing_evaluation_practices_in_philanthropy

A March 2016 piece by Peter Tatian at the Urban Institute details the differences between performance measurement and evaluation and discusses the idea of a performance measurement-evaluation continuum to help organizations and funders identify which activities are most appropriate to undertake.

Tatian, P. (2016, March 15). Performance Measurement to Evaluation. Retrieved from https://www.urban.org/research/publication/performance-measurement-evaluation-0

Appendix C: Glossary of Terms

BASELINE The situation before the initiative. This might be about individuals or the population. You need to have this information to be able to see how far you have progressed towards achieving your outcomes, or to carry out pre/post analysis. (Superu) *Other Terms: Starting point*

BENCHMARK Benchmarking is about comparing your results with others, or to some external standard. To do this you need comparable information, which could be quantitative or qualitative. (Superu) *Other Terms: Standard, Reference Point*

COLLECTIVE IMPACT The commitment of a group of important actors from different sectors to a common agenda for solving a specific social problem. (Kania and Kramer)

CONTINUOUS QUALITY IMPROVEMENT An ongoing cycle of collecting data and using it to make decisions to gradually improve program processes and monitor progress towards stated goals. (Office of Adolescent Health, DIP) *Other Terms: Continuous review, Plan-Do –Study-Act, Quality improvement*

EVALUATION The systematic and objective assessment of a project, program or policy, its design, implementation, and results. Typically rooted in social science research methods, the aim is to determine the relevance and fulfilment of objectives, efficiency, effectiveness, impact, and sustainability. It should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision-making process. Evaluation in some instances involves the definition of appropriate standards, the examination of performance against those standards, an assessment of actual and expected results and the identification of relevant lessons. (Superu)

FOCUS GROUP A group of people selected for their relevance to an evaluation that is engaged by a trained facilitator in a series of discussions designed for sharing insights, ideas, and observations on a topic of concern. (CDC)

IMPACT The change that can credibly be attributed to an intervention. The long-term, big picture change at a population, society or place level. (Superu) *Other Terms: Effect, Contribution to change, Long-term outcome*

INDICATOR A specific, observable, and measurable characteristic or change that shows the progress a program is making toward achieving a specified outcome. (CDC) *Other Terms: Measure, Metric*

INPUTS Resources that go into implementing a program in order to successfully carry out the intervention, such as funds, staff, technology, partners and materials. (Lampkin and Hatry, DIP, CDC) *Other Terms: Resources*

LOGIC MODEL A graphic that shows how a program is intended to work and achieve results, often based on a specific theory of change. A logic model depicts a path toward a goal, typically including inputs, activities, outputs, and outcomes. (Knowlton and Phillips, 2013, as cited by Tatian)

MONITORING The process of consistently measuring over time and reviewing the delivery and results of an initiative so that you can adapt it in the light of findings, or correct it where problems are identified. (Superu) *Other Terms: CQI*

OUTCOME An enduring change in participants' attitudes, emotions, knowledge, behavior, health, or social condition brought about by a program's intentional actions. An intermediate outcome is a more immediate result (e.g., job training participants who obtain full-time employment). A long-term outcome is a more enduring change further removed from the program's direct results (e.g., participants increase economic well-being) (Lampkin and Hatry). *Other Terms: Results, Impacts*

OUTPUTS The direct products of program activities; immediate measures of what the program did. (CDC)

PERFORMANCE METRICS A quantifiable measure an organization uses to track its progress towards carrying out its goals and/or mission. (DIP) *Other Terms: Measure, Indicator*

QUASI-EXPERIMENTAL STUDY A methodology in which research subjects are assigned to treatment and comparison groups typically through some sort of matching strategy or natural experiment that attempts to minimize the differences between the two groups in order to approximate random assignment. (Superu)

RANDOMIZED CONTROL TRIAL (RCT) A process in which participants are assigned to control or treatment groups at random, meaning that all members of the sample must have an equal chance of being selected for the control or treatment groups. (Superu)

RESULT What happens as a consequence of the policy or program. This word is used in many ways in different contexts and should be treated with care – always check how it is being used in each specific case. (Superu) *Other Terms: Outcome, Impact*

SURVEY A data collection method that involves a planned effort to collect needed data from a sample (or a complete census) of the relevant population. The relevant population consists of people or entities affected by the program (or of similar people or entities). (CDC)

TARGET A specific, planned level of a result that you work to achieve within a set timeframe. (DIP)

THEORY OF CHANGE A comprehensive description of how and why a desired change is expected to happen in a particular context. It tries to understand the causal links between what you do and what change occurs – why you believe it will happen. (Superu)

Glossary References

Kania, J. and Kramer, M. (2011). Collective impact. *Stanford Social Innovation Review*, *9*(1), 36-41. Retrieved from https://ssir.org/articles/entry/collective_impact

Lampkin, L.M. and Hatry, H.P. (2003 May 9). Key steps in outcome management. *Urban Institute*. Retrieved from https://www.urban.org/sites/default/files/publication/42736/310776-Key-Steps-in-Outcome-Management.PDF

Superu. (2017 September). Evaluation terms for the social sector: A glossary. *Superu:* Social *Policy Evaluation and Research Unit*. New Zealand. Retrieved from http://superu.govt.nz/sites/default/files/Glossary%20final.pdf

Tatian, P.A. (2016 March 15). Performance measurement to evaluation. *Urban Institute*. Retrieved from https://www.urban.org/research/publication/performance-measurement-evaluation-0

US Centers for Disease Control and Prevention [CDC]. Introduction to program evaluation for public health programs: A self-study guide. *Program Performance and Evaluation Office (PPEO)*. Retrieved from https://www.cdc.gov/eval/guide/glossary/index.htm

US Office of Adolescent Health. (2017). Continuous quality improvement part 1: Basics for pregnancy assistance funds programs. US Department of Health and Human Services. Retrieved from https://www.hhs.gov/ash/oah/sites/default/files/cqi-intro.pdf