Aspen Network of Development Entrepreneurs

The State of Measurement Practice in the SGB Sector

Genevieve Edens & Saurabh Lall
Aspen Network of Development Entrepreneurs

June 2014

Acknowledgements

This report would not have been possible without the support and participation of many organizations and individuals. First, we are grateful to the members of the ANDE Metrics and Research Working Group for identifying the need for this study, providing valuable input into the development of the methodology, as well as peer review. We are also grateful to our external reviewers, Kelly McCarthy and Ellen Carey of the Global Impact Investing Network, and Robert Medina of the Aspen Planning and Evaluation Program, for their thoughtful feedback on early versions of this study. We are also thankful to Mike McCreless of Root Capital and CJ Fonzi of Dalberg Global Development Advisors, for co-authoring the section on our shared vision for the future of measurement.

Finally, we are thankful to the impact investors and capacity development providers that participated in this study, and shared their data and experiences with us. A full list of participating organizations can be found in Appendix I.

Cover photo courtesy of TechnoServe

Table of Contents

EXECUTIVE SUMMARY

The Aspen Network of Development Entrepreneurs (ANDE) is a global network of organizations that propel entrepreneurship in emerging markets. ANDE members provide critical financial, educational, and business support services to small and growing businesses (SGBs) based on the conviction that SGBs will create jobs, stimulate long-term economic growth, and produce social benefits. Helping our members measure their benefits, and assess their impact on the SGBs they support is a core part of ANDE's mission. This report reviews the current state of measurement in two types of organizations that directly support SGBs – impact investors and capacity development providers. Using a mixed-methods approach, we collected data and interviewed over 30 organizations across these two categories, and analyzed key trends in measurement practice.

Impact Investors

We found that the majority of impact investors were doing some type of measurement, mainly through 3 "lenses": Collecting metrics related to both the scale and depth of impact; tracking economic development indicators such as job creation and taxes paid; and specialized sector metrics to benchmark their portfolios. While some impact investors used the findings to track trends and identify patterns, in some cases, they also used the data to assess potential reputational risk, or to track progress against specific targets. Finally, the majority (nearly 80%) of the investors we interviewed mentioned using the findings to report to current funders, as well as to attract new funders.

Capacity Development Providers

The capacity development (CD) providers we interviewed also varied in the way they approached and used measurement. Over half the CD providers surveyed entrepreneurs annually to collect financial performance and job creation metrics, as well as satisfaction with the program. Some organizations focused on measurement as part of the mentoring program itself. Many CD providers use the findings to report to funders, attract new funders, and for communications and marketing. However, we also found a significant percentage of internal use of metrics, to inform operational and strategic decision making.

Emerging Themes

As we move forward, we see a number of emerging themes. Several organizations cite the need to balance and align social metrics with financial performance indicators, to place greater emphasis on transparency and attribution, and to develop more efficient data collection and data management approaches. Our research provides an overview of the current state of measurement in the SGB sector, recognizing the tremendous progress over the past five years. Early efforts were driven by the need for *accountability*, followed by a push for *standardization*. As the state of measurement in the sector continues to evolve, we lay out a vision for the next phase, in which we hope to focus on *value creation*. We hope the analysis and examples, along with our proposal for the future of metrics will provide insight and inspiration for organizations that support SGBs in emerging markets.

I. INTRODUCTION

Understanding the impact of supporting small and growing businesses (SGBs) is at the core of ANDE's mission. Since our launch in 2009, we have implemented various initiatives to support evaluation and performance measurement for the SGB sector. Some of these efforts include the annual Metrics from the Ground Up Conference, support to our members on the adoption of IRIS (the common reporting language for social and environmental performance), and academically rigorous evaluations of SGB impacts through the research initiative. Given the tremendous growth in metrics and evaluation over the past five years, we recognize the need to take stock and understand the current state of measurement in the SGB sector. We believe that understanding why SGB intermediaries choose to measure their performance, which approaches they use, and the challenges they face, will help us develop more effective, rigorous, and useful measurement for the sector.

This report focuses on the state of measurement practice in the SGB sector. We define measurement practice as any mission-related metrics and data collection, analysis and reporting. This measurement may be for internal learning, for reporting to funders, or any other purpose. Measures may be financial, social or environmental, and may be part of a broader theory of change, or simply to track progress towards a target. Ultimately, measurement is about figuring out how well we are doing, and (hopefully) helping us improve.

We focus on the following key questions in this study:

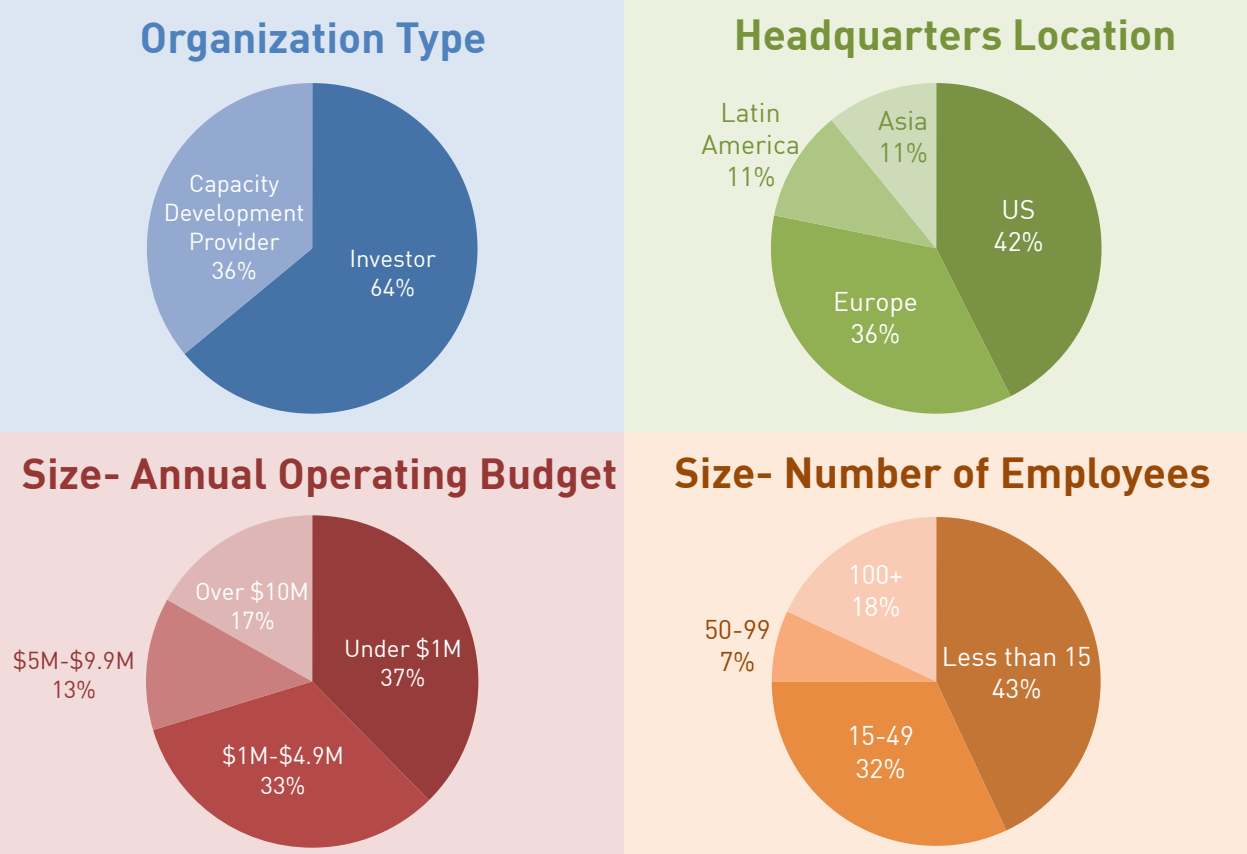
- Why do SGB intermediaries measure their performance and impact?
- What methods, tools and approaches do they use?
- How do they use the findings?
- What are the main challenges they face, with respect to measurement?

This report is primarily aimed at the kinds of organizations discussed in it - intermediaries that provide capital or capacity development to small and growing businesses in developing countries. It aims to provide a snapshot of measurement practices in the SGB sector, identify existing gaps and recommend potential solutions.

II. METHODOLOGY

We used a mixed-methods approach to collect and analyze data, combining hour-long, semi-structured interviews with survey data. Overall, we conducted 33 interviews with staff from 31 organizations. Additionally, we received survey responses from 34 organizations. Questions for the survey and interviews were developed in partnership with ANDE's Metrics & Research Working Group, and based on related studies in the nonprofit sector.[1] The qualitative responses were coded to identify patterns and common elements of evaluation practice across organizations. Given the differences between investors and capacity development providers, we analyze both groups separately, to dig deeper into specific characteristics of their operating and reporting models.

Figure 1: Study Participants



These organizations are very diverse (Figure 1). We do not consider the sample of 34 organizations studied here to be representative of the broader SGB ecosystem, and suggest caution in generalizing results to the broader sector. However, we believe these findings can provide useful insights into the metrics and evaluation activities of the sector, as well as identify areas of future development for the sector.

1 *State of Evaluation 2012, Benchmarking Evaluation in Foundations,* and *Monitoring, Evaluation, and Learning in NGO Advocacy.*

III. TRENDS AMONG INVESTORS

All 20 impact investors in our sample had either developed, or in the case of newly formed funds, were developing, some framework to measure social performance. We identified three "lenses" through which they perceived the impact of their investments, which influenced their choice of measurement approach, as summarized in Table 1.[2]

Table 1: Impact Lenses

Impact Lens	Measurement Approach	Percentage of Investors
Investments create benefits to livelihoods for the **base of the pyramid**, and access to goods and services for disadvantaged communities	**"Theory of Change"** approach that maps how each SGB's activities relate to ultimate impact. Focus on metrics that related to both the scale and depth of impact	59%
Investments create **economic development**, including job creation and value chain effects	Track **economic development metrics** such as job creation and taxes paid. Also, conduct **case studies** to explore the ripple effects of investments.	50%
Investments have a **specific sector focus** such as health, financial inclusion, education, agriculture	Specialized sector metrics to **benchmark the portfolio or calculate the value of social return**. May link outputs to well-defined outcomes based on clear criteria for business models that generate the most impact.	41%

2 A single investor could identify with more than one lens, so the total number is greater than 100%

Why do Impact Investors Measure?

ANDE asked impact investors why they measured impact, to better understand how organizations frame the purpose of their investment in measurement. Among the 18 investors who responded to this open-ended question, the most common answers related to the organizational model.

- **50%** of interviewees pointed out that their Limited Partners (LPs) are investing for a development purpose, so the General Partner (GP) must demonstrate development impact.
- **44%** of investors pointed out that measurement is imperative to their mission; because the fund was created for a social purpose, they must measure the social performance not just the financial returns.

These responses are not necessarily surprising. They mirror two of the core characteristics of impact investors according to the Global Impact Investing Network: intentionality (the intent to generate social and environmental impact) and measurement (commitment to measuring social and environmental performance).

In addition, 40% of investors referenced a motivation to check back on the success of investments. Other motivations that were not as widely cited included: field building, helping SGBs increase their impact, and to improve decision making.

The fact that decision-making ranked low in frequency was also not surprising: anecdotally, one trend in the sector is that impact investors are just beginning to integrate social performance data into investment decisions and portfolio management. We hope to see this integration will increase, as more investors see the value in informing business decisions with social and environmental performance data.

"The majority of our investors are DFIs, and they're interested in investing for impact."

"It's really core – in no sense is it an option. It's part of [the organization's] heartbeat."

"We are an impact investor, so we need to demonstrate impact."

What Measurement Approaches do they use?

Most of the investors we interviewed invest long-term capital, either equity or quasi-equity, and so develop long term relationships with their investees. Nearly two-thirds of investors measure social and environmental metrics during the due-diligence phase, and about half will measure over the course of a longer relationship with the SGB. Typically, during the due diligence stage investors assess the SGB's current performance and potential for future impact. Then during the life of the investment, or over the course of a longer relationship, investors track performance and occasionally assess outcomes.

1. **Measure at the Due Diligence Stage - 67%**
   - *Criteria Screen:* 30% of investors assessed potential investments against an established set of criteria, based on the fund's vision for impact (which may be sector-specific).
   - *Estimate Outcomes:* 39% of investors estimate potential outcomes prior to making an investment, typically by reviewing the evidence base to better understand the potential of a particular business model for generating impact in a target community. A smaller percentage reported developing formal estimates. For example, they may calculate how much a smallholder farmer's income could increase through access to improved seed, or farmers' return on investment if they purchase the company's product.
   - *Theory of Change:* About one-fifth of investors work with each potential investee to map the connection between the enterprise's activities and the desired outcomes, along with the assumptions that support the theory of change.
   - *Field Visits:* Finally, 3 investors also made field visits as part of due diligence or during the investment period, and spoke with company staff, clients, or suppliers to anecdotally verify the potential outcomes on the lives of ultimate beneficiaries.

2. **Portfolio Approach - 44%**
   - *Scorecard or Rating System:* 39% of investors reported using a scorecard or standard rating system to benchmark individual investments (and potential investments), and to ensure that they were maintaining a positive social return at the portfolio level. See Appendix III for a detailed description of these frameworks.
   - *Linking Financial Compensation to Social Performance:* Two fund managers link their financial compensation through carried interest to social performance of their portfolio. One fund manager targeted a quantified social return for the portfolio as a whole. (See page 9 for detailed descriptions of these approaches.)

3. **Post-Investment Measurement: Outputs vs. Outcomes**[3]
   - *Measuring Outputs:* 61% of investors focused on tracking outputs over the course of their relationship with the SGB, and did not measure outcomes.
   - *Deep-dive Studies:* 39% of interviewees tracked outputs but also conducted deep-dive studies to examine outcomes with a selection of the portfolio. These studies included qualitative case studies, quasi-experimental studies, or quantitative household surveys.
   - *Encouraging SGBs to Measure:* Finally, 20% of investors consider the SGBs themselves to be ultimately responsible for measuring outcomes or tracking social performance data at the level of the beneficiary. They encourage, incentivize, and help build capacity so enterprises can integrate measurement themselves. In one case, for example, SGB measurement systems are part of the criteria screen.

Fifteen investors described their output measurement framework. Of these, six use a standard set of metrics across their whole portfolio, four customize metrics for each investment, while

3 Outputs are the products, services, or facilities that result from an organization's activities. Outcomes are the changes, benefits, learning, or other effects that happen as a result of what the organization offers or provides.

five use a combination of standard and customized metrics. 75% of the investors we interviewed used IRIS-aligned metrics to track enterprise-level outputs.

About 60% of investors track metrics related to scale (for example, number of clients reached or number of units sold), while over 40% also included metrics related to quality or depth of impact (for example, agent churn or payments to producers). In a few cases, investors also tracked environmental and social governance (ESG) metrics - policies related to socially and environmentally responsible management.

How do they use the findings?

Understanding how impact investors actually use the results from these practices is essential in developing more effective methods and tools. Of the 12 investors that spoke about use, about 40% mentioned using the data to **track trends and identify patterns**, for specific businesses as well as across their portfolios. In some of these cases, tracking social performance was also linked to financial performance reporting. For example, one investor is interested in providing access to new services in rural, underserved communities, and tracks sales revenue alongside the proportion of clients in rural areas. The investor can then track overall growth and financial success of the company alongside the mission-driven target market expansion.

Four investors mentioned that they use the data to **assess potential reputational risk**. Typically, these were metrics around policies and practices, for example environmental management practices or human resource policies, and were more common in agriculture-focused investors. About one-third used the data to **track progress against specific social performance or scaling targets** for an investment.

Finally, in a few cases, investors highlighted the importance of using the social performance data to **understand outcomes**. Two investors calculated the value of their social return, adapting the Social Return on Investment (SROI) methodology. Two investors also tried to estimate the impact of their own investment into the SGB, by looking at changes as a result of new business lines or expansion into new markets. These investors saw their impact as a driver of scale, so outcomes could be measured by the growth of the company.

In addition to use, fourteen investors also discussed how the findings added value to their operations. The majority (nearly 80%) mentioned **reporting back to current LPs** and **attracting new funders**. In many cases, the adoption of measurement practices signals a commitment to social impact to potential investors. Over half the investors also reported that measurement helps them **manage their relationship with SGBs**, to help them improve performance and align with the investors' expectations. Finally, about half the investors also said the **SGB may find value** from the reporting practice itself, as a learning experience.

Challenges

The biggest challenge that appeared to be common to about 40% of the investors was the lack of stability and resources for measurement that their investees face. Most early stage ventures are still experimenting with their business models, and find it challenging to measure their social performance in a formal manner. Some investors also cited challenges with their investees' internal record keeping system (or lack thereof), and said that many companies saw these reporting requirements as a burden.

Social Performance Based Success Fee

In traditional private equity, fund managers structure their earnings in two ways: a management fee that covers the cost of investing, and a success fee that incentivizes performance. This success fee, called carried interest, is paid to the fund manager once they've hit a certain level of return, or hurdle rate. Once the pre-determined hurdle rate is met, the fund manager receives a portion of the profits over that amount, usually around 20%. This income structure incentivizes fund managers to improve the fund's performance, since they directly benefit from higher returns. The private equity funds that invest for a social purpose as well as financial return use this same compensation structure. But a few pioneering impact investors have noted the possible incentive mismatch with financial return-based compensation structures, and have sought to incentivize strong social performance as well through carried interest.

Vox Capital

Vox Capital is an impact-focused venture capital fund that invests in early stage enterprises that benefit the base of the pyramid in Brazil. They state that "if the fund delivers beyond its benchmark financial return rate, we only have access to the full carry compensation if we also reach a certain level of impact. If we do not reach the minimum expected social impact level, our team is entitled to only half of the carry. At the same time, if the financial return is below its target, we are not entitled to any success fee, regardless of our social impact results, avoiding another potential conflict of interest: managers earning money, even when investors lose it." [4]

In order to assess whether the fund has met its social benchmark, Vox Capital rates its portfolio on social and environmental standards using GIIRS. Vox chose GIIRS because it is the pioneer in standardizing impact ratings, and was one of the first impact funds to adopt the rating system. The rating combines the fund management approach together with individual company assessments that are aggregated to reflect the average social performance. The portfolio as a whole must reach a certain standard of social performance, agreed upon in advance by Vox and its LPs. While the hurdle is set at the overall GIIRS score, Vox places more focus on the Community and Impact Business Model metrics based on their BOP focus.

Insitor Management

Insitor Management is a fund manager based in Cambodia that invests seed capital in social businesses in India and Southeast Asia. The fund seeks to fill a critical gap for low-income households in the region, and supports enterprises that provide new access to goods and services for the base of the pyramid. The fund's investors are aligned with this mission; given the early-stage, social business environment, all the three funds that Insitor manages target below-market rate returns. As Associate Marco Aletti points out, "[our investors are] giving us patient capital – so they want to make sure social impact has been achieved." This need for accountability has resulted in an income structure where carried interest is received by the fund manager only if social targets have been achieved alongside financial profit. These social targets are set on a company by company basis rather than at the portfolio level because of their investees' diversity. Insitor works closely with each firm to determine social performance metrics and the corresponding target. Thus, incentives are aligned along the whole chain of capital, from enterprise to GP to LP.

4 Daniel Izzo, "Aligning Interests in Impact Investing," Stanford Social Innovation Review Blog. Accessed from: http://www.ssireview.org/blog/entry/aligning_interests_in_impact_investing

So you want to measure outcomes?

For many ANDE members, the small and growing businesses they work with are part of a means to an end. Their ultimate goal might be poverty alleviation, economic growth, or equitable access to a critical service. SGBs are a mechanism through which members achieve those goals. The impact framework for these organizations references the clients, suppliers, employees, or distributors of SGBs. Among the investors we spoke with, there were three main approaches to evaluation that moved measurement beyond the output level of the SGB itself.

Qualitative Case Studies

The members who took this approach tended to select one company to investigate each year, and produce a detailed report. One staff person (or independent contractor or academic) leads the project, typically traveling to the site of operations to spend about one week, and up to one month, collecting information about the company through interviews and surveys. Data sources include employees at all skill levels, suppliers, clients, and occasionally competitors. For example, the study may include interviews with current employees to assess whether they earn more at the study company than they did at their previous job. It may calculate the company's multiplier effect on the local economy, and assess the "indirect jobs" created in the value chain.

Stakeholder Surveys

In a few cases, members that wanted to connect with the ultimate beneficiaries of their investment, (employees, suppliers, or clients) reached out directly through surveys. These surveys did not include a control group, so rather than trying to understand the impact (relative to what would have happened anyway, without the SGB), they focused on assessing beneficiary perceptions. This kind of study may also assess changes over time against a benchmark, but without including a control group. Members noted that these types of studies were typically more cost-effective to manage in-house, and still produce the kinds of information they needed.

Experimental or Quasi-Experimental Studies

This type of study was least common among the organizations that we interviewed. Research of this kind requires locating a group to serve as the counterfactual, so essentially doubles the investment of resources into the investigation. In these cases, the member partnered with an academic, who designed the study. The member might then outsource the management of the research entirely, or might provide project management support as data is being collected. Analysis is conducted by the academic partner.

Table 2. Outcome Measurement on a Shoestring Budget

	Field Work	Human Resources	Est. Cost	Key Questions
Qualitative Case Studies	1 week	1 researcher	$5,000 - 10,000	• What are the local economic impacts of this firm?
Stakeholder Surveys	1 week	Lead researcher, survey enumerators	$25,000 - $50,000	• What are the poverty levels of the firm's stakeholders? • Do stakeholders see value in the firm's operations?
Experimental or Quasi-experimental studies	Long-term (6 months – 3 years)	Lead researcher, survey enumerators	$100,000+	• What is the effect of this firm's activity on a group of stakeholders, compared to a similar group that did not receive any benefits from the firm?

IV. TRENDS AMONG CAPACITY DEVELOPMENT PROVIDERS

"Measurement helps us improve our programs and services for participants. It helps us keep in mind why we're here, and what we're doing this for."

ANDE interviewed 11 Capacity Development (CD) Providers that work with entrepreneurs to develop skills, build management capacity, and raise growth capital. Two of the organizations were large, international nonprofits that employ thousands of people across multiple regions, and manage a range of programs focused on SGBs. The rest were smaller organizations that typically focus on a particular SGB segment (for example, women entrepreneurs or social enterprises). The CD providers we spoke with offered a wide range of services, such as cohort-based mentoring and online mentoring programs.

Why do Capacity Development Providers Measure?

The majority of capacity development providers reported measuring for internal purposes: improving programs, checking success, and learning. Some organizations also cited field building as a motivation. Interestingly, CD providers did not commonly cite reporting to funders as a motivation.

"Measurement helps us understand whether we are doing what we set out to do. If we don't see benefits then we can ask ourselves if we're taking the right approach, whether we should change our project design, or whether something has changed in the external environment."

What Measurement Approaches do they use?

The capacity development providers that we interviewed provide a diverse set of services to entrepreneurs, making it difficult to draw general conclusions. Capacity development providers typically engage directly with SGBs for shorter periods of time, compared to investors. Once the entrepreneur leaves the program, CD providers generally try to continue data collection for multiple years, to get a sense for the program's longer-term impact on the business. Although the majority of organizations we spoke with were aligned in their motivations to measure for internal learning, the measurement approaches were quite varied.

- Over half the organizations s**urveyed entrepreneurs annually**, asking questions about firm-level measures such as financial performance and job creation metrics. About 40% also asked about the **entrepreneurs' satisfaction** with the program, providing them with relevant feedback about program quality.
- Some CD providers focused on measurement as **part of the mentoring program itself.** For example, helping each entrepreneur to develop a theory of change or define key performance metrics. In some cases, the CD providers even set specific social targets for the entrepreneurs to track their performance against.

- Only the two large international NGOs conducted **household level surveys** to understand outcomes at the level of the SGBs' clients, suppliers and employees. They also mentioned a set of core performance metrics that were tracked across all projects, in addition to project-specific metrics and more in-depth research projects.

How do they use the findings?

We also asked CD providers how the results were ultimately used. We found high mentions of external data use, including reporting to existing funders, attracting new funders, and communications and marketing. Consistent with the stated motivations, we also found considerable internal use of the findings. Over 60% highlighted their use of performance data in operational decision making, and about half said that it was also used to inform longer-term, strategic decisions.

For example, one organization highlighted that the strategic shift it made recently towards an increased focus ***"on the most promising members of our portfolio, those with growth potential"*** was based on the data they had been collecting. ***"It took several years to make that change, and the data was very important."***

Another organization highlighted the direct operational use for performance data collection. ***"We have achieved a high understanding of... who is in the program, who is doing well, and who needs more support. We have that grasp of what's happening in the program, and it helps us improve."***

Challenges

Capacity development providers cited challenges similar to those faced by impact investors (for example: the resource and capacity constraints of the SGBs). Additionally, these organizations mentioned their own unique resource constraints.

Three CD providers highlighted the challenge of communicating the social impact of diverse business models, especially for donors not familiar with the field. When CD providers work with very diverse businesses in multiple sectors, the "breadth of different kinds of impact that companies have and finding ways to aggregate" to tell one compelling story becomes challenging. Similarly, for larger, decentralized organizations the diversity in each office's priorities can make it difficult to standardize measurement while still allowing for regional diversity.

NESsT: A portfolio approach to capacity development

NESsT's mission is to solve critical social problems through the development of sustainable social enterprises. It does this first through incubation support that includes tailored technical assistance and mentoring. The highest performing enterprises that are ready to start the scaling process then receive financial investment in the form of long-term loans.

NESsT takes an exceptionally rigorous approach to performance measurement during the incubation stage, and has developed a portfolio approach to assess not only the performance of an individual enterprise but to benchmark the entire cohort of enterprises across the organization. NESsT sees their emphasis on measurement as a benefit not just internally, but to the social enterprises themselves and the field as a whole.

NESsT's portfolio is tightly clustered around 3 social impact objectives (Table 3), which allows them to create standard output metrics across their portfolio. They recognize that financial sustainability is critical to achieving these social targets, so put equal weight on goals for financial milestones, and tracking financial performance.

To help these start-up social enterprises track their performance effectively, NESsT developed a "Performance Management Tool," designed to not just measure progress, but actually help enterprises scale. The PMT sets social enterprise and organizational goals from the business plan and due diligence process, and measures progress with corresponding quantitative and qualitative indicators. It is a tool for social enterprise managers to set goals and targets and regularly monitor performance against them. It is their way to report progress to NESsT. It is also a tool for NESsT portfolio managers to outline needs that will be addressed during the specific implementation year and identify potential weaknesses.

The enterprise fills out the PMT three times a year; NESsT compiles this performance data into their Salesforce.com database, and conducts an internal portfolio assessment based on 6 categories for benchmarking. It assigns a score from 1 -3 for each category:

- Skills and knowledge increased
- Active use of the PMT
- Consolidated team and leadership
- Potential for scaling
- NESsT fit
- NESsT value added exists

Enterprises whose aggregate score is above 15 are considered ready for the "scaling phase," those between 11-14 are recommended to "stay in incubation for another year" and those below 11 are considered for "immediate or planned exit within a year." Final decisions take into account both their performance in one of the three impact areas, as well as the benchmarking assessment.

"We are looking for the potential for sustainable and scalable social enterprises when selecting applicants for our planning stage; we are working to strengthen the business models that will be able to deliver social impact in incubation stage; we are evaluating track record and potential to increase social impact when deciding about scaling social enterprises."

Table 3. NESsT Impact Goals

Impact Area	Target Population	Impact Target	Sample Specific Output Metrics
Labor Inclusion	People with disabilities, at-risk youth, ethnic minorities, and other populations excluded from the labor market.	600 beneficiaries per enterprise	Number of jobs provided to marginalized communities. • Number of people trained and placed in jobs or whose employability has increased through training. • Number of people placed in the open job market.
Sustainable Income	Artisans, small-scale farmers, and other low-income and marginalized producers	750 beneficiaries per enterprise and 30% increase in income	Number of people receiving income. • Number of people who gain skills and knowledge to improve their products/services to increase sales. • Percent increase in individuals' income.
Affordable Technology	Low-income communities	Quality of life improved for 20,000 per enterprise	Number of people benefiting from new technologies. • Long term positive impact of technologies on target population. • Number and type of technologies offered and new technologies developed. • Geographical reach of new technologies and their expansion.

Table 4. NESsT's Portfolio-Level Metrics

Enterprise Performance	Targets and key indicators of the business performance, including revenue, costs, profits, etc.
Social Impact	Targets and indicators for the social impact, including beneficiary numbers and benefits and systemic change effects
Institutional Development	Targets and indicators relating to the institutional capacity, including governance, HR, and stakeholder relations
Financial Sustainability	Targets and indicators relating to the diversity of financial resources, including asset building and revenue streams

Aliança Empreendedora: An alternative framework to assess entrepreneurs' performance

Aliança Empreendedora is a non-profit organization that supports the design, testing, and implementation of inclusive business models, and entrepreneurship for low-income communities. They work with low-income entrepreneurs and cooperative groups of artisans, supporting them to grow their businesses. Initially, Aliança Empreendedora had been tracking performance based on the financial success of the businesses that the entrepreneurs started. They often saw, however, that financial performance metrics did not capture the complete picture of their program's impact. An entrepreneur might start a business that struggled, but as the entrepreneur developed this flawed business she was learning a huge amount. In 2 or 3 years, Aliança staff found, she might start a new business that was much more successful. Aliança Emprendeedora decided that rather than the financial performance of a particular business, they should focus on the entrepreneurial development of each program participant.

Aliança staff had read about the theory of effectuation from academic Saras Sarasvathy. The theory resonated strongly with some of the patterns they had seen within their own program: successful entrepreneurs start with their "own means," and imagine goals based on those internal resources. Then the entrepreneur builds support for his or her idea by enlisting others from a broader network. This happens in an iterative process, until the entrepreneur has developed a successful business, including a product with committed customers and other stakeholders.

With this realization, Aliança decided to radically shift the focus of their performance measurement from financial metrics to metrics related to effectuation (with the addition of income, as the sole remnant of their previous framework) (Table 5).

The shift towards an effectuation framework created a more complex set of assessment tools, and Aliança worked over several months to simplify their questions and the way in which they were asked. Ultimately, the organization moved towards self-assessment forms that entrepreneurs fill out, using pictographs instead of written questions. This makes the surveys accessible to entrepreneurs who may have low literacy levels.

The entrepreneurs fill out the assessment every 6 months. Because the tool is in the early stages of implementation, at this point it is mostly useful as an assessment for program staff to identify areas of strength and weakness, and target their interaction with clients to those areas that need more support. Eventually, Aliança plans to use the data to assess the pace at which entrepreneurs are improving along all dimensions of their framework, as well as the performance of specific interventions or staff approaches.

Table 5. Effectuation Framework for Entrepreneurial Development

Skills: Who I am	• Self-knowledge and self-confidence • Vision of the future • Hands on approach • Will to learn new things / skills • Persistence • Personal relationship
Knowledge & Experience: What I know	• Financial controls • Pricing • Time management • View of the markets • People management • Formalization • Production management • Marketing • Sales and negotiation
Network: People I know	• Contacts with clients • Participation in networks and events • Contacts with other entrepreneurs • Contacts with potential partners and investors • Contacts with suppliers • Contacts with inspiring entrepreneurs
Income: How much I earn	• In the last 6 months has your income ... (increased, stayed the same, decreased) • Can you maintain yourself with your own income? • There was something you wanted and were able to buy/make with this income? • Did you fulfill a dream through your business? What?

TechnoServe: A comprehensive corporate measurement framework

TechnoServe is a nonprofit that works with enterprising people in the developing world to build competitive farms, businesses, and industries. To do this, TechnoServe acts as a catalyst to strengthen market systems by developing the capacity of market actors, promoting market connections, and improving the business-enabling environment. TechnoServe operates in more than 30 countries and employs around 1,400 people, but U.S. headquarters staff is fairly small. Headquarters provides support to country offices on centralized processes such as finance, HR, operations, marketing, and business development. The country offices typically manage multiple projects and have their own staff to lead and implement these projects.

While TechnoServe has consistently valued measurable results at both the project level and organizational level, in 2011 the organization decided that improvements were needed to its corporate measurement approach. In the previous framework, definitions of indicators were not consistently applied across geographies, results were not disaggregated by gender and did not consistently reflect incremental and attributable results associated with TechnoServe's work. In addition, quality control was inconsistent and field staff felt burdened by the reporting requirements. As a result, TechnoServe was missing out on an opportunity to use measurement effectively as a management tool.

The New Corporate Measurement Framework

Senior leadership across the organization wanted a system that provided critical input into three main functions: 1) Strategy, planning and management decision making; 2) Learning about differences in interventions, context, and what works across the portfolio of TechnoServe's work; and 3) Internal and external communications. \

The refreshed corporate measurement system was developed to help measure results against TechnoServe's overarching mission. It was to provide a common standard across diverse projects for learning, aggregation and comparison. This system would complement the more in-depth individual project monitoring and evaluation (M&E) systems. Over the course of a year, a cross-functional global working group researched metrics being used by similar organizations and donors, gathered feedback on the current TechnoServe system, and drafted a set of core indicators that would be collected across all of the organization's projects.

Table 6. TechnoServe Corporate Performance Metrics

Category	Key Question	Metric	Notes
1. Project Participants	Who is better off because of our work?	Number of farmers adopting promoted practices	Disaggregate by gender; where feasible, count only those individuals who changed their behavior, not all participants
		Number of businesses assisted	
		Number of jobs created	
2. Financial Benefits	How much better off are the participants?	Incremental attributable farmer revenues	Where feasible, compare participant performance with a counterfactual
		Incremental attributable business revenues	
		Incremental attributable wages	
3. Efficiency	How efficient are we at creating benefits?	Return on (TechnoServe) Investment ("ROTI")	Incremental attributable benefits / Project costs
4. Sustainability	Are the benefits likely to be sustained once the project is over?	Total value of debt and equity finance mobilized by project participants during project implementation	Sustainability only determined years after the project. Metric is used as a proxy as it reflects a source of capital for farmer/entrepreneur to invest and reflects external confidence in sustainability

The system had to be flexible, but ensure high quality data. Toward that goal, several levels of quality control are built into the system. Data, including notes about assumptions, are first entered online by project or M&E staff based in the field. The data is then reviewed and approved by the Project Manager, Country Director, Regional Director, and finally COO. Once approved, summary and project detail reports are accessible to all staff.

Rollout of the new corporate measurement system

Designing a metrics framework is the (relatively) easy part; the difficult step is convincing users to actually collect, enter, and use the data. TechnoServe's new system was designed to be more relevant, meaningful, and useful to staff in the field based on their feedback. For example, in the new system, each project submits data annually but can choose the schedule based on other reporting requirements. Teams are required to report only on the subset of headline indicators that are relevant to their project.

To get staff on board with the new system, TechnoServe piloted the new indicators with a subset of countries and updated the process based on their feedback. The person who led the rollout had been at TechnoServe for more than five years and had strong connections across the organization. By demonstrating the benefits of the system, she was able to convert some of the early detractors into champions and successfully implement the system in partnership with staff across the organization. By the time 2012 reporting was closed out, all projects had contributed data into the system.

Putting the Data to Use

One of the complaints about the previous system was that staff saw it as a burden rather than a potential management tool. This new framework, with consistent metrics and strong quality control, was designed to be useful for internal decision-making. With results from just one year in the new system, TechnoServe's global leadership was centering strategic planning and operational discussions around aggregated project data. Regional and country leaders examined data across projects to understand how TechnoServe creates impact and which potential investments should be prioritized.

At the project level, teams can now benchmark their performance with similar projects in different geographies. In one case, a team accessed the database to look for similar projects; they contacted the other implementing teams to learn about the approaches that could improve the efficiency and effectiveness of their own project. As TechnoServe staff expands their use of the database, the organization is planning to invest further in reporting functionality to help facilitate data-driven decision making, transparency, and learning.

V. NUTS AND BOLTS

Resourcing Measurement

In addition to in-depth interviews, 33 organizations participated in a survey that asked specifically about the percentage of their operating budget and staff time that went towards measurement in the last fiscal year.[5]

These organizations are diverse, with staff size ranging from 2 to 5,000, and annual operating budgets from $200,000 to $74 million. Revenue sources are equally diverse: on average, 40% of revenue comes from grants, 40% from fee for service, and 20% from other streams. About 70% of the organizations interviewed received 40% or more of their total revenue from private foundations or individuals, 25% received 40% or more of their revenue from DFIs or donors, and 8% received a majority of revenue from corporations. Operating budgets were relatively "unrestricted," meaning they were not tied to specific project and could be used for any purpose, with a median 61% of the budget unrestricted.

Within this landscape of diverse revenue models and funding sources, there was an equally diverse range of spending (both staff time and budget allocation) on measurement (Table 7, Table 8). We asked respondents about the staff who participate in measurement, both the number of full time and as a percentage of individuals' time. Allocation of total staff resources to measurement ranged from 0% - 30%. The total spending on measurement (calculating staff time and consultant fees) as a percentage of the operating budget ranged from 0% - 25%. The majority of organizations used internal resources rather than hiring outside experts; only 4 out of 21 organizations spent more than 45% of their measurement spending on consultants.

Table 7. Resource Allocation – All Organizations

Median number of staff (FTEs) engaged in measurement	1.5
Percentage of organizations with at least 1 full time measurement-focused staff person	48%
Median percentage of total staff focused on measurement	5%
Median annual spending on measurement	$50,000
Median measurement spending as percentage of total budget	2.2%

5 Of the 33 organizations that responded to this survey, 28 also completed an interview, so the samples are overlapping but not identical.

Table 8. Resource Allocation by Organization Size* and Type

Organization Size	Larger	Smaller
Capacity Development Providers		
Median Spending	$ 71,980.00	$ 12,980.00
Median Percentage of Budget	2.91%	1.02%
Min - Max Percentage of Budget	0.7% - 4%	0% - 2.3%
N	6	6
Median FTE	3.9	0.2
Median Percentage of Staff	5.6%	1.0%
Min - Max Percentage of Staff	0.4% - 13%	0% - 4%
N	8	6
Investors		
Median Spending	$ 125,000.00	$ 20,000.00
Median Percentage of Budget	2.2%	8.1%
Min - Max Percentage of Budget	0.6% - 10%	0.7% - 25%
N	7	6
Median FTE	1.8	0.3
Median Percentage of Staff	5.4%	6.0%
Min - Max Percentage of Staff	0.1% - 14%	3% - 30%
N	8	11

*We split the sample at the median staff size, and labeled all organizations with more than 24 staff as "larger" and those with 24 and fewer staff as "smaller."

Table 9. Benchmarks for SGB Sector Spending on Measurement

Given this study's small sample size, direct comparisons to other stakeholders is not possible. But benchmarking can give sense for where ANDE members might aim in terms of resource allocation. External benchmarks are based off of similar studies conducted in different industries.

	Spending levels less than 2.5%	Spending levels 2.5 % - less than 5%	Spending levels 5% and higher
International Development Projects	51%	49% *2.5% and higher*	
US-based Nonprofits	73% *less than 5%*		27% *5% and higher*
ANDE Members	60% *less than 2.5%*	20% *2.5% - less than 5%*	20% *5% and higher*

Sources: State of Evaluation 2012, The State of Development Evaluation

For the organizations that participated in the interview process in addition to completing the survey, we looked more closely at differences in resource allocation as it related to their measurement practices. We found a statistical tendency for organizations that mentioned an internal use of the data to spend more on measurement (p=.06, n=27). Those with internal use allocated on average 7% of total human resources compared to 3% for organizations that did not use the data internally. We suspect that the more data is utilized internally, the more valuable it becomes to the core operations, and organizations subsequently increase spending on measurement.

Organizational Structure, Data Collection, and Data Management

Half of the capacity development providers we spoke with have a separate department focused on measurement, while just over 10% of investors have a separate department (Table 10). Those organizations without a separate department integrate measurement into other roles within the organization, or in some cases hire a single staff person focused on measurement.

Table 10. Structure of the Measurement Function

	Investor (n=16)	CDP (n=10)
Measurement is integrated into other departments	88%	50%
A separate measurement department exists	13%	50%
At least 1 full-time measurement staff person	44%	70%
N	16	10

The majority of staff that lead the measurement effort at their organization report directly to the CEO. Other organizational structures include reporting to the COO, the Impact or Innovation Officer, communications, or program management staff.

Data is collected most commonly on an annual basis, although investors may collect data more frequently than capacity development providers (Table 11).

Table 11. Data Collection Frequency by Organization Type

	Investors (n=14)	Capacity Development Providers (n=8)
Annually	6	4
Monthly	3	
Quarterly	2	1
Semi-annually	2	
Other /Irregular based on program timelines	1	3

Typically, both investors and capacity development providers decentralize data collection responsibilities (Table 12). The person or department with the closest relationship to the SGB is in charge of collecting data from the SGB. In a few cases, an organization had previously structured data collection so that one person reached out to the entire portfolio of companies, but they found they could increase the response rate if the main relationship manager sent out the data collection request. The measurement process also benefited from translation into local languages, building trust, and integrating into other operational activities through this strategy. On the other hand, for those organizations with a centralized data collection process (with their metrics officer working closely with each portfolio company to validate the data), the metrics staff person had a broader picture of the entire portfolio's performance data.

Table 12. Data Collection Responsibilities

Investors (n=14)	
Decentralized -- Investment Manager/Officer	86%
Centralized -- Portfolio Analyst	14%
Capacity Development Providers (n=7)	
Decentralized -- Program Staff	71%
Centralized -- Measurement Staff	29%

Over half of the organizations that described their relationships with SGBS reported that they develop a formal agreement, such as an MOU with each SGB/entrepreneur that specifies data reporting expectations. Of these 9 organizations that do use formal agreements, seven are investors and only two capacity development providers.

Once the data has been gathered from each SGB, the organizations we interviewed typically store the aggregate data in Excel, a Salesforce-based database, or in another cloud-based system (Table 13).

Table 13. Type of Technology

Investors (n=12)	
Excel	7
Cloud system developed in-house	2
eFront	2
PULSE / B Analytics *	2
Capacity Development Providers (n=7)	
Salesforce	4
Cloud system developed in house	2
PULSE / B Analytics *	1

**In late 2013, the Salesforce-based performance tracking system PULSE was integrated into the B Analytics and GIIRS platform managed by B Lab*

The data is pulled into reports and dashboards and shown to the organization's key stakeholders, in particular, the board (Table 14).

Table 14. Primary Audience (n=12)

All Staff	50%
Board	50%
Senior Management	33%
Funders	33%
Investment Committee	25%
SGBs	17%

VI. EMERGING THEMES

Financial Growth, Scale, and Social Impact

Organizations' objectives and ultimate goals vary a great deal within the SGB sector. Some intermediaries support growth-oriented businesses regardless of sector because they see the value in the economic growth and job creation that they believe those firms' success will bring. Others support growth-oriented businesses within very specific sectors, or those that adhere to a particular business model with the potential to alleviate poverty. Finally, some organizations support social enterprises regardless of sector; they believe in the power of business solutions to a range of social and environmental challenges, and see their work as supporting the evidence for that approach as a whole.

Many organizations across this spectrum recognize a tension between financial growth and social impact. In some cases, the organization's primary focus was financial viability for the SGB. These organizations believed that businesses would not have the chance to generate social impact without financial viability, and emphasized growth.

In other cases, the tension depended on the business model. ***"It's always easier with companies that have the clearest impact stories,"*** one interviewee remarked. For example, a company that sells solar lanterns can easily estimate their impact through the number of lanterns sold; their sales growth directly links with the scale of their social performance, or the number of households with access to clean lighting. But some business models may drive more intangible, systemic change rather than easily quantifiable metrics.

Tensions arose even in cases where the business model aligns financial and social performance. For example, in cases where revenues are closely linked to the number of beneficiaries (customers or suppliers), the question of pricing still creates challenges. A company that sells seed at a reduced price, for example, might be reducing its margin to optimize its social performance. Similarly, a company that buys product from suppliers at the base of the pyramid may reduce its purchase price, increasing its margin at the expense of social performance. ***"This is a strategic decision,"*** another interviewee pointed out. ***"There needs to be a discussion with the entrepreneur and management team. And for that you need data."***

Analysis at the level of the business model, and understanding the firm's relationship with "impact" is a direction that the field needs to expand. Data around social performance by business model is one area that has not yet emerged. Clarity and transparency in this area will create better educated LPs, more effective fund managers, and stronger social impact.

Attribution & Transparency

A few organizations that we spoke with brought up the challenge of attribution, especially in the context of field building. Often, more than one fund manager co-invests in a single firm; in reporting back to investors and the public, the fund managers ***"are essentially double counting"*** their reach. As the field grows, some see the need for a transparent methodology to account for each investment's appropriate share of the impact. Transparency seems to be the first

step in this direction. Rather than focusing on the external communications end, some organizations advocated for sharing data and methodology between fund managers, and developing a consistent approach to attribution. Creating this transparency in the field would be a first step towards scale, one organization pointed out. While at the level of the fund manager, self-reporting and unique methodologies are necessary to avoid the risk of over-simplification, third-party verification and controls are also necessary to truly scale impact investing.

Automating Data

Many organizations pointed out the data collection process itself as "labor-intensive," "inefficient," and "challenging." This was especially true for Capacity Development Providers, whose programs often require periodic surveys with entrepreneurs who they no longer directly support. It was also a challenge that was mentioned in the context of the SGBs themselves. They often did not have strong systems in place to track social performance data internally, so sending it to other stakeholders was challenging. Embedding data collection processes into the work of programs and businesses is something that will greatly ease the burden of data collection, and lead to more nuanced, high quality data availability. Automating the collection process itself will allow organizations to spend more time using the data rather than gathering it.

VII. THE FUTURE OF MEASUREMENT: Introducing a New Vision for Shared Metrics

Mike McCreless, Root Capital; CJ Fonzi, Dalberg Global Development Advisors; Genevieve Edens & Saurabh Lall, ANDE

There has been tremendous progress in measurement over the past five years, with the establishment of common standards and the recognition that metrics are an essential part of operations for intermediaries that support SGBs. We are encouraged by the fact that every organization we spoke to had either already implemented some measurement strategy, or were in the process of doing so. The growing adoption of the IRIS standards also reflects the continuing maturity of the sector, as it aligns around a common language to measure its performance. We have seen the focus of measurement in the sector shift from *accountability* (Metrics 1.0), to *standardization* (Metrics 2.0).

However, while SGB intermediaries can continue on their own unique paths of performance monitoring and impact evaluation, as well as report on common metrics year after year, our research leads to the question: where do we go from here?

We believe that the opportunities created by impact measurement and evaluation have yet to be fully realized, and that there is another step-change – Metrics 3.0 – that we should aim to reach within three years.

A Vision for Metrics 3.0

We believe that the next phase of metrics will shift the emphasis from *accountability* (Metrics 1.0) and *standardization* (Metrics 2.0) to *value creation* (Metrics 3.0).

For metrics and evaluations to create value for us, individually and collectively, we must do two things:

1. **Put our metrics to work by integrating impact metrics with financial and operational ones.** Integrated metrics can help organizations develop better products and services, improve resource allocation, and build more efficient and impactful businesses.

2. **Implement targeted, actionable evaluations that are useful to multiple stakeholders, and fit with collective learning agendas.** Such evaluations will build on existing knowledge, break down big questions into manageable, answerable pieces, and put the answers back together to inform strategic decision-making for enterprises and for the sector at large.

The below chart lays out our vision for how Metrics 3.0 will drive business value and impact at the organizational level, and help us to achieve our shared goals and mission as a sector.

Figure 2: Four Windows to Create Value

Value Creation	Organizational-Level: *Data-driven decision making*	Ecosystem-Level: *Growing the playing field*
Measure: *Putting Your Metrics to Work*	1 **Use what you know** Integrate social, environmental, financial, and operational metrics to enhance enterprise impact and profitability	2 **Share what you know** Attract resources by demonstrating collective results, better allocate scarce resources, and share costs of data collection and innovation
Evaluate: *Learn what you need to know*	3 **Learn what you need to know** Evaluation that serves multiple purposes: test assumptions; demonstrate impact; gather customer feedback; target specific unknowns to build a better product, service, or delivery model	4 **Build a learning agenda** Build an actionable evidence base, enable practitioners to access and repurpose what is known, and prioritize evaluation efforts to fill gaps

1. Organization-Level Measurement

Most organizations track social and environmental metrics as a separate function; the data is stored separately, analyzed by dedicated staff, and reported on in its own publications. Our vision for Metrics 3.0 shifts isolated impact metrics to integrated financial, operational, and impact metrics.

Integrating financial and impact metrics is a challenging task. Existing forms of cost-benefit analysis, developed for policy-makers and service delivery non-profits, don't quite fit the new breed of social entrepreneurs, impact investors, and mission-driven businesses. The challenge for metrics professionals is to become sufficiently fluent in operational and financial metrics to be able to integrate them with impact metrics, and to create and advocate for this integrated approach with the leaders of their enterprises.

2. Ecosystem-Level Measurement

We believe that the potential of aggregated performance data has not been realized. At a sector level, Metrics 3.0 continues the unfinished work of Metrics 2.0 by encouraging demand-led aggregation of impact data that creates value by demonstrating the scale or reach of a group of organizations, or by enabling capital providers to compare organizations' impacts and efficiency more effectively. More broadly, we believe that aggregated data will only be useful when it is being collected with a clear purpose.

3. Organization-Level Evaluation

Evaluations are expensive. Using the highest standards of rigor, it's possible to spend more evaluating an investment than the value of capital deployed. But we believe that organizations can creatively, less expensively, conduct evaluations that create more value for additional stakeholders, including the people who are participating in the evaluation (e.g., 'being studied'). Hard data about consumers and producers living at the bottom of pyramid is scarce and valuable to other actors, like upstream companies in the value chain, NGOs serving the same community, and donors and investors. Each data collection exercise represents an opportunity to create value in multiple ways:

- By testing assumptions about our impact, and hopefully, demonstrating impact
- By refining our products and services
- By channeling feedback from low-income consumers or producers to other stakeholders engaged in the community (who may in turn be willing to share the cost of the evaluation)

4. Ecosystem-Level Evaluation

Evaluation, like most forms of knowledge creation, is a public good. Not every organization will conduct evaluations, but every organization could benefit from evaluations done by others. What's more, organizations working together can build a base of evidence about 'what works' that no single organization would be able to build alone.

We envision that in Metrics 3.0, sector-wide initiatives will catalogue existing information and findings and identify gaps. In a virtuous circle, this will drive individual organizations to develop evaluations that are "repurposable" – that is, useful to multiple stakeholders.

Conclusion

The Metrics 3.0 framework we described will enable organizations to maximize the value of impact metrics by integrating them with financial and operational metrics, to inform day-to-day decision making, as well as longer term strategic planning. It also proposes how multiple enterprises in a sector can collaborate to demonstrate collective scale, channel resources to the most impactful and efficient activities, and start build a base of evidence about 'what works' around a shared learning agenda. We look forward to working with ANDE members to take action in each of the four windows of opportunity, and moving to the next phase of measurement in the SGB sector.

This section originally appeared as a post in the Stanford Social Innovation Review blog in June 2014.

Appendix I: Participating Organizations

Organization	Representative(s) Interviewed	Job Title(s)
Accion	Sonia Arenaza	Director, Channels and Technology
Acumen	Tom Adams	Director of Impact
Agora Partnerships	Dorrit Lowson	Chief Operating Officer
Bamboo Finance	Ximena Escobar de Nogales	Head of Social Performance Management
Calvert Foundation	Kevin Fanfoni & Andrew Parucci	Portfolio Analytics Officer, Marketing Officer
Cherie Blair Foundation	Allison Kahn	Mentoring Programme Coordinator
d.o.b. foundation	Hedwig Siewertsen	Managing Director
Dasra	Arjav Chakravarti	Impact Assessment Manager
Global Social Benefit Institute	Andrew Lieberman	New Programs Director
Grameen Foundation	Stephanie Simpson	Portfolio Analyst
HUB	Hinnerk Hansen & Alyona Asyamova	Managing Director, Global Impact Coordinator
I&P	Elodie Nocquet	Financial & ESG Officer
Impact Finance	Benjamin Firmensch	Investor Relations Partner
Insitor Management	Marco Aletti	Associate
LGT Venture Philanthropy	Tom Kagerer	Chief Operating Officer
Media Development Investment Fund	Peter Whitehead	Director of Communications
Mercy Corps	Jon Kurtz	Director for Research and Learning
NESsT	Eva Varga	Director of Portfolio Performance
NOTS Foundation	Sandra Kingma	Managing Director
Oxfam	Hugo Sintes & Alan Doran	Enterprise Development Programme Manager, Business and Finance Adivser
Pipa	Dhaval Chada	Founding Partner
responsAbility	Christian Etzensperger	Senior Research Analyst
Root Capital	Mike McCreless	Director of Strategy and Impact
Small Enterprise Assistance Fund (SEAF)	Mildred Callear & Tanja Atanasova	Executive Vice President, Portfolio Manager
TechnoServe	Jennifer Golden & Krisila	Senior Project Manager, Operations; Senior Director, Program Services
TriLinc Global	Marni Hodder & Kathryn Haugen	Chief Impact Officer, Impact Analyst
Vital Voices	Marguerite Berger	Vice President for Impact Evaluation, and Research
Vox Capital	Gilberto Ribeira	Partner
Voxtra	Gaute Ellingsen	Investment Manager
Youth Business International	Laura Grisby & Annie Barber	Monitoring Evaluation and Learning Manager

Appendix II. Scorecards

The majority of organizations that we spoke with track a set of core metrics across all of the SGBs they work with. In a few cases, these core metrics were actively used to compare and benchmark across the whole portfolio by combining those core metrics into scorecards and ratings. Below are four rating methods that members described.

A. The Global Impact Investing Rating System (GIIRS)

Overview

GIIRS is a product of B Analytics, a customizable platform for measuring, benchmarking and reporting on impact. GIIRS Ratings are based off of a company and fund's performance on the B Impact Assessment (BIA). The BIA framework is applicable not just to small and growing businesses in emerging markets, but to all businesses, from start ups to large companies on a global scale. There are over 70 versions of the BIA that contain different sets of questions depending on the business size, geography and industry. The BIA is aligned with the IRIS taxonomy, and includes more than 30 core IRIS metrics, and references the IRIS Glossary Terms throughout the BIA.

Framework

The fourth version of the BIA was launched in January 2014, and reflected significant modifications from version 3 to capture feedback from the entrepreneur and investor communities. The BIA framework includes two main components: the Operational rating and the Impact Business Model rating. The BIA provides an overall score out of 200 possible points. The GIIRS Rating Methodology takes the data submitted through the BIA and provides both an Operations rating (1-5 star scale) and an Impact Business Model Rating (Bronze, Silver, Gold, Platinum, or N/A).

Operations

The first half of the BIA focuses on the operational impacts of a business. Questions cover Governance, Workers, Community & Environment, and are tailored to the relevant version of the BIA. The concepts covered here are analogous to typical corporate ESG issues. For the Operations section of the BIA, in coordination with its Standards Advisory Councils, B Lab has aggregated the best operational practices research established in the field. Rather than creating from scratch a set of performance standards, B Lab is leveraging the thought-leadership of a broad range of organizations, particularly groups that focus deeply on particular aspects of business practice (Fair-Labor Organization and International Labor Organization, to give examples from the Worker section) to build a holistic positive impact structure.

Companies see more detailed questions depending on what they do. For example, a manufacturing company is more likely to have greater environmental impact than a service company; so manufacturers are asked specific questions related to environmental impact, and the Operational Environmental section is weighted more heavily for manufacturers. Similarly, a small company has fewer workers, and therefore fewer worker-related questions as it would be expected to have less complex worker practices and policies than a larger business.

Impact Business Model (IBM)

The section of the B Impact Assessment that has been an area of most research for version 4 is the Impact Business Model (IBM). In particular, the impact investor community has pushed B Lab to create a comprehensive framework for measuring business model impact that is more refined and requires greater specificity in reporting. On version 4, B Lab has introduced a uniform, multi-factor approach to each IBM captured in the B Impact Assessment. The framework covers:

1. **Intent of Business** – To qualify for an IBM, the company must demonstrate that the intervention is part of the core intent of the business (rather than an ancillary or unintended impact). For example, to qualify as a Workforce Development/Job Creation business model, a company must hire a certain percent of workers from chronically underemployed communities.

2. **Intensity of Intervention** – Each IBM has questions about how intensive an intervention is (i.e. - how many hours of training and topics covered through workforce development programs) and what percent of the company's business focuses on that intervention (i.e. - through amount of revenues from the sale of a socially or environmentally focused product or service).

3. **Efficacy** – Measuring efficacy or outcomes is the most challenging aspect of impact measurement and also the area of greatest interest. For some IBMs, we have enough data to ask about specific outcomes (i.e. - in workforce development, does a company track the workers it has trained after they "graduate" – beyond existing employment – and have those workers attained gainful employment in the medium and long-term). In other areas, particularly around outcomes associated with a specific product or service, B Lab asks more general questions regarding whether a company has defined desired outcomes, created targets for meeting those outcomes, measured against those targets and have had success in meeting the targets. This approach represents a first step forward to encourage enterprises to think about outcomes, link key outputs to desired outcomes in their business operations, and search whether existing research exists to indicate whether desired outcomes can be achieved.

4. **Beneficiaries served** – This factor focuses on whether an in-need beneficiary is being served, covering employees, suppliers, customers, distributors and the environment (ie for workforce development, the BIA measures whether chronically underemployed individuals have been hired).

5. **Reach** – All IBMs include one or more metrics that measure the scale of intervention. For many, this is a measurement of individuals, households, businesses or communities served, but also covers environmental metrics like tons of CO2 off-set.

6. **Sector innovation** – This factor focuses on identifying those game-changing business models that will revolutionize a sector or create follow-on innovation up or down the value chain. The intent is to capture how a business model innovative, scalable and replicable. We'll gather data from businesses on version 4 to better understand how entrepreneurs define sector innovation, with the goal of adding more nuanced questions on this topic in future versions.

7. **Industry Data** – Leveraging the great work of others, including the Social Performance Task Force within the Microfinance Industry, the Global Alliance of Banking Values in financial institutions, and IRIS' sector working groups, we've integrated industry specific metrics to generate greater industry specific benchmark data.

B. Grameen Foundation – MOTIV Scorecard

Overview

Grameen Foundation's capital markets program invests in Social Enterprises and Microfinance Institutions that improve the economic livelihoods of the poor in rural and difficult to reach markets in line with Grameen Foundation's overall charitable mission to help connect the world's poor to their potential. To ensure alignment between the investment strategy and its charitable mission, Grameen Foundation developed a scorecard called MOTIV. This scorecard is used during desk analysis when initially evaluating company impact, during on-site due diligence for corroboration of impact, and tracked on an ongoing basis as part the portfolio review process. A version of MOTIV was developed for Social Enterprises with emphasis on the investment proposition from the perspective of the poor, and a similar but separate version for MFIs with added focus on treatment of clients and innovative product offerings.

Framework

MOTIV uses a similar framework to Grameen Foundation's models for assessing company and country risk. A company's overall impact & alignment score assesses 40 indicators (both subjective and quantifiable) across five overall categories, ranging from the organization's social intent to a particular value proposition from the perspective of the poor. Scores are rated on a scale of 1 (lowest) to 5 (highest) with specific guidance for assessing each indicator built into the model. For example, when assessing the relative affordability for a poor entrepreneur to engage in a business opportunity, a score of "1" for least affordable is generated if the investment cost for the entrepreneur is greater than 20% of per capita GDP, while a score of "5" is assigned for most affordable if the entry cost is less than 2.5%.

Scores for each indicator are automatically calculated based on input values, and aggregated for a total MOTIV score. Scores are mapped to a 100 point index which can be used to compare results across the portfolio ranging from "Highest Impact/Strategic Fit" at one end to "Reputation Risk" at the other.

Although MOTIV was designed to align with GF's specific social mission and strategic directive, the model could be readily calibrated to match the social priorities of other organizations addressing poverty by simply adjusting the weighting of each indicator as desired, or further customized (with more effort) by replacing certain indicators and modifying the underlying scoring guidelines or formulas.

Figure 3. Sample MOTIV Scorecard

Company:	Sample Social Enterprise	Country: Region:		Date: 12/31/13
INDEX	CATEGORY	CLASSIFICATION	SCORE / CODE	WEIGHTING
3.00	MOTIV Score (Social Impact & GF Mission Alignment - Social Enterprise)	Strong Impact	70	100.0%
		CLASSIFICATION	SCORE / CODE	% of Total Score
3.10	Mission & Organization Commit.	Moderate Impact	66	17.5%
3.11	Focus on Poverty Reduction (Intent)	Strong Poverty Focus	4	3.0%
3.12	Social Alignment of Investors & Stakeholders	Good Alignment	4	2.5%
3.13	Social Alignment of Mgmt. & Employees	Good Alignment	4	2.0%
3.14	Local Ownership	Negligible Ownership	2	2.0%
3.15	Balanced Return Expectations	Undefined	2	2.0%
3.16	Dividend Payout Ratio to External Stakeholders	n/a	3	2.0%
3.17	Treatment of Employees (Benefits, Opportunity)	Above Average	4	2.0%
3.18	Pay Disparity	30.0 x	3	2.0%
3.20	Outreach	Strong Impact	72	22.5%
3.21	Size of Entrepreneur or End Beneficiary Client Base	Small Client Base	2	6.0%
3.22	Total % of New Clients Below the Poverty Line	45.0%	4	6.0%
3.23	Growth % of # of Clients	80.0%	5	1.4%
3.24	Inclusion of Poorest as Employees & Beneficiaries	Strong Inclusion	4	2.5%
3.25	Ongoing Connectivity With the Poor	High Connectivity	5	2.5%
3.26	Use of Technology to Maximize Outreach	Moderate Utilization	3	1.4%
3.27	% of Rural Clients	75.0%	5	1.4%
3.28	% of Women Clients	50.0%	3	1.4%
3.30	Transformational Impact	Moderate Impact	57	17.5%
3.31	Scalability	Moderate Scale Potential	3	4.5%
3.32	Replicability	Low Replicability	2	3.0%
3.33	# of FTE Local Employees	35	4	3.0%
3.34	Access to New Markets / Business Opportunity	Moderate Linkage	3	2.0%
3.35	Access to Critical / Actionable Info	n/a	1	2.0%
3.36	Environmental Impact	Neutral Impact	3	1.5%
3.37	Other Positive Externalities	Above Avg. Impact	4	1.5%
3.40	Investment Proposition for the Poor	Highest Impact	81	25.0%
3.41	Incremental Income Potential	100.0%	5	5.5%
3.42	Entry Cost Affordability for the Poor (Notional)	$250	3	3.0%
3.43	Entry Cost (As % of Per Capita GDP)	7.1%	4	3.0%
3.44	Financing Availability	Reasonably Available	4	3.0%
3.45	Time to Break-Even on Investment (months)	3	5	3.0%
3.46	IRR% to the Client	150.0%	5	3.0%
3.47	Diversification of Assets or Income Flows	Moderately Diversified	3	1.5%
3.48	Reliability of Income Flows	Fairly Reliable	3	1.5%
3.49	Frequency of Income Flows	Occasional	2	0.8%
3.50	Liquidation Value of Investment / Asset for Client	0.0%	1	0.8%
3.50	Verification & Measurement	Moderate Impact	62	17.5%
3.51	Implementation of PPI or Similar Product	Piloting	4	4.5%
3.52	Frequency & Depth of Survey Sampling	Anecdotal	2	2.5%
3.53	Staff Ability & Resource Dedication for M&E	Below Avg. Capacity	2	2.5%
3.54	SPM Reporting & Measurement of Appropriate KPIs	Good SPM Metrics/Reports	5	2.5%
3.55	Analysis of PPI & SPM Applied to Operations	Not Applied to Date	1	2.5%
3.56	Third Party Social Rating / Certification	Above Avg. Rating	4	2.0%
3.57	Recency of Last Social Rating / Certification	22 mo.	3	0.9%

C. Root Capital

Description

Root Capital lends and provides financial management training to small and growing agricultural businesses in Latin America and sub-Saharan Africa. They operate with a social and environmental purpose, and unlike the majority of ANDE members, manage a large portfolio of short-term loans, typically working with 200+ clients each year. In 2011, Root Capital revised and expanded the Social and Environmental Scorecards that they use to evaluate prospective clients. The data captured in the scorecard focus on outputs and operational practices of the enterprise rather than outcomes, which are more difficult to standardize and accurately observe in the loan application process.

Framework

The scale and output metrics that Root Capital tracks are IRIS aligned, but the metrics tracking operational practices are outside the IRIS framework.

Category	Practice
Scale	Number of Producers Reached by the Business
	Number of Women Reached by the Business
	Number of Employees
	Number of Female Employees
	Wages and Salaries Paid
	Volume Purchased from Producers per Crop Type
	Hectares Under Cultivation by Affiliated Producers
	Payments to Producers
	Enterprise Revenues
Additionality and Context	Access to Finance of the Business
	Vulnerable Contexts (ie violence, natural disasters)
	Poverty Level of Producers Affiliated with the Business
	Enterprise Growth
Higher and More Stable Incomes for Producers	Enterprise Pays Higher Price to Producers than Other Local Buyers
	Enterprise Raises Local Market Prices for All Producers
	Enterprise Wages and Benefits to Employees
	Enterprise Enhances Producer Productivity in Existing Activities
	Enterprise Creates New Income-Generating Activities
	Enterprise Provides Credit to Producers
Producer and Employee Rights and Treatment	Occupational Health and Safety
	Rights of Association of Employees and Producers
Community-Level Practices	Enterprise provides Public Goods (e.g. health, educational programs)
	Gender Inclusive Practices
Environmental Practices	Environmental Management Systems
	Land Use and Biodiversity Conservation
	Agrochemical Usage
	Soil Erosion Control and Fertility Management
	Water Usage and Wastewater Management
	Solid Waste Management
	Energy Sources and Efficiency of Usage

Each enterprise receives an impact rating from C to AAA. Ratings of "C" disqualify a loan from financing, thereby, helping loan officers exclude SGBs that don't meet Root Capital's criteria. By contrast, loans are designated "Strong Social Impact" if 1) the enterprise and its associated producers have significant unmet needs, and 2) enterprise increases the level and / or stability of farmer incomes. Strong environmental performers are those that receive an A or above in each of the seven environmental performance categories, indicating that the enterprise and/or its producer suppliers are practicing agriculture in a manner that sustains and/or enhances farm-level environmental health. In 2012, Root Capital went through an exercise to ensure calibration across staff members scoring enterprises, and verify the quality of data by cross-checking with third-party sources. In 2013, Root Capital undertook an extensive cost allocation exercise to estimate cost and profit per loan, and to create tools and dashboards that integrate loan-level and portfolio-level impact metrics with financial metrics, for use by loan officers and senior managers.

Root Capital's Social and Environmental Scorecards, along with a methodology guide, can be found here: http://info.rootcapital.org/social-and-environmental-due-diligence

D. Impact Finance

Impact Finance – Kharmax

Description

Impact Finance provides debt financing to cooperatives, small businesses, and microfinance institutions in varying positions along the value chain, including production, processing, and distribution, both in emerging and developed economies. Impact Finance developed a proprietary rating system that evaluates an enterprise's impact from a value chain perspective. Some metrics are standards for all companies, while others are sector or project specific. Impact Finance cusomizes the metrics based on a particular commodity, for example gold and agricultural crops have different factors that are included in the overall score.

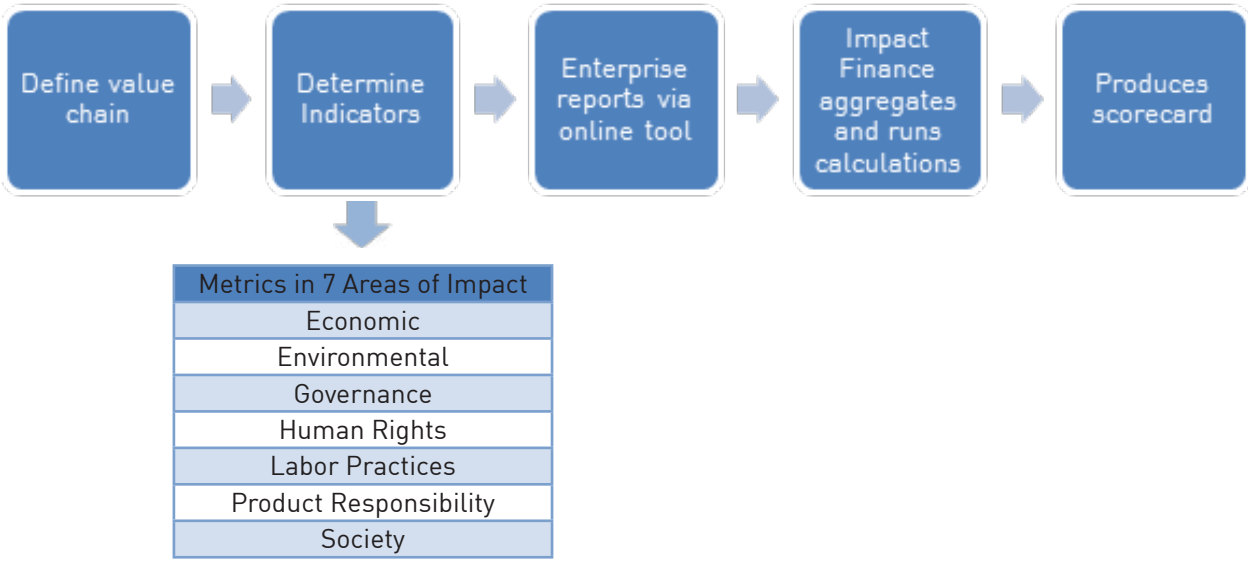
Framework

For each company, the impact of the overall value chain is calculated through 7 categories, each composed of a number of aspects. Some of those aspects include output metrics for the enterprise, for example the number of clients or suppliers. These metrics are used to calculate indicators that range from 0 to 3, by adding information about the quality as well as scale of impact. For example the enterprise may calculate the average monthly net income for its employees, which is compared to local minimum wage to build an indicator from 0 to 3. Other indicators are closed-ended questions where companies assess themselves ticking one of four answers. For example "Profit participation scheme" can be between 0 (No scheme for distributing profit to personnel/members) and 3 (There is a comprehensive scheme in place with the entire staff/members included).

Result

The end product for each company is a scorecard that is available online. Impact Finance works not only with its own portfolio, but also other fund managers and corporations that source from these value chains, to rate the enterprise. For fund managers and corporations that work with multiple companies, scores of each category are aggregated into a Fund Score. Self-assessments by companies can be validated by an on-site audit that leads to the generation of an audit report.

Figure 4. Impact Finance Framework & Methodology



Category	Aspect
Economic	Sustainability of the Business Model
	Distribution of wealth inside the actors
	Fairness of the Value Chain
Environmental	Biodiversity conservation
	Energy consumption
	Use of natural resources
	Effluents and waste
	Water consumption
Governance	Social mission and commitment
	Representation of stakeholders in the Governance
	Organization, independence and qualification of the governing body
Human Rights	Child labor
	Right of association
	Indigenous rights
Labor Practices	Employee health and safety
	Labor relations
	Access to training
Product responsibility	Product quality
	Customer health
	Communication on product conditions
Society	Community services
	Anti-corruption systems
	Involvement in sector promotion or public policy lobbying

Appendix III. Literature Review

While the research on measurement in the SGB sector is relatively limited, there has been considerable interest in studying the topic in the nonprofit and for-profit sectors. A number of studies have looked at monitoring and evaluation practices in the UK, US, and internationally. We summarize some key points from these studies that are relevant to the SGB sector, to help us benchmark our focus and approach compared to different but related sectors.

Nonprofits in the UK

According to a 2012 study, the majority of nonprofits in the UK (75%) evaluate their activities in some form (N=550). Evaluation in the UK is primarily driven by funders that require reporting, with 52% of respondents stating that they conduct evaluation primarily to report to funders, compared to 22% that report to the board. Only 5% say that improving services is the primary motivating factor, while 4% say it is to see how well they are performing, which suggests that evaluation is mainly used for reporting and accountability, and may not provide information that is relevant for internal planning or decision making. The study also notes that organization size is an important factor, and many smaller nonprofits do not conduct evaluation activities. In the larger organizations, leadership appears to be a strong driver of evaluation, while medium sized organizations are mainly motivated by funder reporting requirements. While the study found a generally positive perception of evaluation, they noted that about 60% of their sample stated that they do not get enough funding for evaluation, and 21% say that monitoring and evaluation actually interferes with their work. Finally, the study found that the stated motivations for conducting evaluation may not always match up with the benefits that they receive from it.[6]

Nonprofits in the US

In the United States, the State of Evaluation report examined evaluation practices in nonprofits in 2012 and 2010. In 2012, they found that 94% of the organizations in their sample conducted evaluation (up from 85% in 2010). They found that about 70% of the organizations spent under 5% of their annual budget on monitoring and evaluation, and 38% received no support from funders for these activities. Organizational size was also a factor, with larger nonprofits being more likely to conduct evaluation. Interestingly, the primary audience for evaluation findings changed from funders in 2010 to the CEO in 2012, with 74% stating that the findings were directly reported to internal organizational leadership. While all nonprofits state that they use the findings (to plan and revise programs, and report to funders), the study found that organizations faced several challenges in implementing evaluation, including limited staff and resources. Evaluation was rated as one of the lowest organizational priorities in both years.[7]

Foundations

Recent research also suggests a shift in how funders approach evaluation. A recent study of US foundations found that 50% of 31 US based foundations had increased support for evaluation

6 Ogain et al. (2012)

7 State of Evaluation 2010, 2012

in the past year.[8] Researchers noted a shift in the focus of evaluation, from "did it work" to "what did we learn to improve our programs" in the approaches that foundations employed. Foundations were connecting programs and strategy to evaluation more closely, pushing it beyond a purely accountability and compliance driven function, and 75% of the respondents had the evaluation director reporting directly to the CEO.

For-profit Performance Measurement

While traditional businesses do not typically conduct impact evaluations, there is a long history of using performance measurement systems that combine financial and non-financial metrics, such as the Balanced Scorecard (Kaplan & Norton, 1996), and the Performance Pyramid (Neely, 2005). While these systems were originally designed as systems to support decision making, they have now evolved to also improve operational efficiency, strategic planning and accounting.[9] However, organizational size appears to be an important factor in the for-profit sector as well, with many researchers finding that the adoption of performance measurement systems in small and medium sized enterprises has been extremely limited. They find that existing frameworks are mainly structured for the needs and capabilities of larger firms, and do not fit the informational needs and resource constraints of smaller firms. [10]

Overall, it is evident that there are significant challenges in performance measurement and evaluation in both the nonprofit and for-profit sectors. The literature suggests that for-profits tend to use it primarily for organizational learning, and that it is largely motivated by internal reasons (such as decision making, strategic planning), while non-profits tend to be motivated by external reasons (such as reporting to funders). Organizational size appears to be a key factor that affects the adoption of evaluation in all cases, with smaller organizations finding it difficult to implement evaluation practices. Finally, we see a shift in foundations' perspectives on evaluation, from a purely accountability driven approach to a learning oriented agenda.

8 Coffman et al., 2013

9 The Balanced Scorecard, in particular, is widely used, and has also been adapted for use in the nonprofit sector.

10 Tatichi et al. (2010); Henri (2006); Grafton, et al. (2010).

Appendix IV. References

Bititci, U., Garengo, P., Dörfler, V., & Nudurupati, S. (2012). Performance measurement: Challenges for tomorrow. *International Journal of Management Reviews*, 14(3), 305-327.

Carman, J. G. (2011). Understanding evaluation in nonprofit organizations. *Public Performance & Management Review*, 34(3), 350; 350-377; 377.

Coffman, J., Beer, T., Patrizi, P., & Thompson, E. H. (2013). Benchmarking Evaluation in Foundations: Do We Know What We Are Doing?. *Foundation Review*, 5(2).

Ebrahim, A. (2005). Accountability myopia: Losing sight of organizational learning. *Nonprofit and voluntary sector quarterly*, 34(1), 56-87.

Ebrahim, A. (2003). Making sense of accountability: Conceptual perspectives for northern and southern nonprofits. *Nonprofit Management and Leadership*,14(2), 191-212.

Ebrahim, A. (2002). Information struggles: The role of information in the reproduction of NGO-funder relationships. *Nonprofit and Voluntary Sector Quarterly*, 31(1), 84-114.

Ebrahim, A. S., & Rangan, V. K. (2010). The limits of nonprofit impact: A contingency framework for measuring social performance. *Harvard Business School General Management Unit Working Paper*, (10-099), 10-099.

Garengo, P., Biazzo, S., & Bititci, U. S. (2005). Performance measurement systems in SMEs: a review for a research agenda. *International journal of management reviews*, 7(1), 25-47.

Innovation Network (2010). *State of Evaluation 2010: Evaluation Practice and Capacity in the Nonprofit Sector*. Innovation Network Inc.

Innovation Network (2012). *State of Evaluation 2012: Evaluation Practice and Capacity in the Nonprofit Sector*. Innovation Network Inc.

Coe, J. & Majot, J. (2013). *Monitoring, Evaluation and Learning in NGO Advocacy: Findings from Comparative Policy Advocacy MEL Review Project*. Oxfam & Overseas Development Institute

Henri, J. (2006). Management control systems and strategy: A resource-based perspective. *Accounting, Organizations and Society*, 31(6), 529-558.

Grafton, J., Lillis, A. M., & Widener, S. K. (2010). The role of performance measurement and evaluation in building organizational capabilities and performance. *Accounting, Organizations and Society*, 35(7), 689-706.

Aspen Network of Development Entrepreneurs