Spring 5-31-2013

Exploring the History and Challenges of Monitoring and Evaluation in International Nongovernmental Organizations: Complemented by Intern Experience at Save the Children USA

Scott Toscano
SIT Graduate Institute

Recommended Citation

EXPLORING THE HISTORY AND CHALLENGES OF MONITORING AND EVALUATION IN INTERNATIONAL NONGOVERNMENTAL ORGANIZATIONS: COMPLEMENTED BY INTERN EXPERIENCE AT SAVE THE CHILDREN USA

Scott Louis Toscano

PIM 71

A Capstone Paper submitted in partial fulfillment of the requirements for a

Master of Intercultural Service, Leadership and Management at

SIT Graduate Institute in Brattleboro, Vermont, USA.

May 31, 2013

Advisor: Professor Marla Solomon

Consent to Use of Capstone

I hereby grant permission for World Learning to publish my Capstone on its websites and in any of its digital/electronic collections, and to reproduce and transmit my CAPSTONE ELECTRONICALLY. I understand that World Learning's websites and digital collections are publicly available via the Internet. I agree that World Learning is NOT responsible for any unauthorized use of my Capstone by any third party who might access it on the Internet or otherwise.

Author's Signature: _____ Date: _____

Acknowledgements

I have so much privilege and opportunity in my life. I am lucky beyond any means and forever grateful to my mother and father. It is with your spirit that I go forward every day and try to spread the joy I have in my life, not to change the world, but to change somebody's world.

Jessica, you inspire me. Your laugh motivates me and your quiet hum brings me to a serene place. Our journey is only just beginning. We will make many memories ensuring to 'leave only footprints and take only photographs.'

I would not have accomplished what I have without a lot of help along the way. I have many caring friends and mentors in my life. You stuck with me through some tough times. You might not think of it like this but you have made all the difference in my life.

My experience in Peace Corps set the stage for this capstone. During my time in Azerbaijan, I partnered with Azerbaijani program managers in Save the Children's Kürdəmir field office on several assignments. Their charisma and dedication is so great we are still friends to this day! After that office closed down, I collaborated with a former staff member to pioneer a sport cycling program using USAID funds. Before I even knew what it was, I was monitoring a program in an international development context.

My advisor, Marla Solomon, provided guidance for my research and writing so I could cull meaning from mounds of data. You helped me to ask the right questions.

Save the Children provided excellent internship opportunities that made this practicum possible. The feedback, comments and edits suggested by my internship supervisors have been invaluable and help make this capstone paper more accurate and interesting. Thank you all so much for your encouragement, support and understanding throughout this process.

TABLE OF CONTENTS

List of Abbreviations

APM	Agency Performance Measurement
CLC	Course-Linked Capstone
CO	Country Office
CP/HIV	Business team for Child Protection and HIV-AIDS (SCUS)
DECD	Business team for Education and Childhood Development (SCUS)
DHL	Business team for Hunger and Livelihoods (SCUS)
DHN	Business team for Health and Nutrition (SCUS)
DHR	Business team for Humanitarian Response (SCUS)
GI	Global Indicator
IE	Impact Evaluation
INGO	International Nongovernmental Organization
KM	Knowledge Management
M&E	Monitoring and Evaluation
MELWG	Monitoring, Evaluation and Learning Working Group
MOS	Management Operating Standard
NGO	Nongovernmental Organization
OECD-DAC	Organization for Economic Co-operation and Development, Development Assistance Committee
PPA	Department of Public Policy and Advocacy
SC	Save the Children
SCUS	US-based Save the Children (independent organization)
SCI	Save the Children International
SMT	Senior Management Team
SOP	Standard Operating Procedure
UN	United Nations
USP	Business team for United States Programs (SCUS)

ABSTRACT

This paper examines the monitoring and evaluation (M&E) system of a large international non-profit organization, Save the Children. An internship was used to support the central M&E unit and work on projects related to its needs. Projects included cataloguing evaluation reports from the past two years and evaluation plans for the upcoming year, interviewing key informants about technical resource utilization and validating evidence about known gaps in Save the Children's M&E system. A review of evaluation literature showing the evolution of the development evaluation practice is followed by a description of the mechanisms in Save the Children's M&E system and the analysis and findings of the research.

The analysis showed that Save the Children's evaluation policies and procedures were thoroughly documented and grounded in international standards. However, research confirmed what M&E personnel believed to be the case: that there is a limited skill level for country level personnel to adequately follow the procedures. Furthermore, there was not uniform storage of evaluation reports. Evaluation plan summaries were catalogued as expected but those products were of limited use. At the time technical resources and tools were not organized well and were not generally accessible by country office level personnel who need them the most. Recommendations include adding administrative personnel and expanding the functionality of the central M&E unit to develop metrics to rate the M&E system that would increase executive oversight of the M&E system. This could incorporate an audit of the M&E system using a peer-review assessment method developed by the Organization for Economic Co-operation and Development.

Introduction

Save the Children's mission is "to inspire breakthroughs in the way the world treats children and to achieve immediate and lasting change in their lives" (Save the Children, n.d.d, Who We Are). Breakthroughs do not happen every day; world-changing breakthroughs do not necessarily happen every week or month or year. Likewise, lasting change can only be seen throughout time. Save the Children believes it takes commitment to accountability and measuring results long-term to reach its goal.

Accountability and learning are the key purposes for conducting program evaluations. Evaluations are a systematic assessment typically examining program design, implementation or results compared to program objectives or another set of criteria. They document outcomes, assess effects and elicit understanding about what worked. The expectation is that the lessons learned through this process will lead to innovations that achieve important changes. So that learning can occur, evaluations are supposed to highlight program successes and failures plus the intended and unintended consequences of implementing the intervention.

Many donors find it challenging to balance the tensions between funding either evaluation and learning or life-saving programs and services. Save the Children recognizes that "programs have not been consistent in integrating evaluations in program design across all sectors" and "…proposes the institutionalizing of an Evaluation and Learning Fund to close this gap" (Save the Children, 4/30/12, p.1). It is imperative for all the mechanisms of an M&E system work in concert to generate the best quality evidence from program interventions. The evidence generated by its evaluations can be used to influence national policies and scale up interventions to impact the lives of more boys and girls around the world.

Large and small nonprofit organizations should be monitoring and evaluating their programs. Each organization tells its story and makes claims using various tools that measure program outcomes and organizational efficiency. It is not necessarily the particular measurement methods utilized that count for an organization's success. Rather it is the way an organization systematizes its efforts and builds on lessons learned. This paper is a Course-Linked Capstone (CLC) for SIT Graduate Institute that is designed to link the theories discussed in its Program Monitoring and Evaluation course to the practice of M&E. This paper will explore the history and challenges of M&E within the context of international development by citing specific examples from an internship experience with Save the Children US. Through this research NGO's can learn about the process of assessing an M&E system, as well as international standards for M&E systems.

Need for this Research and CLC Capstone

The contemporary climate of M&E within which Save the Children operates is dominated by pressure to get results to make the case for more or extended funding. Save the Children explains, "[t]he current political and funding environment continues to stress the importance of demonstrating results and increased accountability. Our donors, charity raters and the public are demanding more evidence to prove near term and longer term impact of our work" (Save the Children, 4/30/12, p.1). Save the Children plus numerous other diverse, mission-driven NGOs perform development interventions locally and globally based on donor interests and guided by their own values. Cooperation and partnership among these aid organizations necessitates international standards for evaluation to direct the practice of M&E and increase accountability to stakeholders, above all beneficiaries. The OECD and its members cultivate standards for international development management, including principles for M&E and M&E

systems. The OECD published *Principles for Evaluation of Development Assistance* (1991) which includes a section on "the management and institutional arrangements of the evaluation system within development agencies" (OECD DAC, 2010, p.5). It is challenging to consistently and systematically apply these core principles throughout an organization, over many years, and under many constraints from both the internal and external environment.

Save the Children recently announced its strategic plan to "invest in impact for children" with a special objective to "invest in innovation and evaluation to measure our impact for children" (Save the Children, n.d.a, p.1). See Appendix A for an illustration of the Save the Children 2013-2015 Strategic Framework. Save the Children examined its M&E system and determined there were "missed opportunities for learning due to the fact that evaluations were not planned and not conducted or they were not of good quality…for the evaluation results to be used" (Save the Children, April 2012, p.1). I began an internship in the central M&E unit to get hands-on experience to learn about the real-life challenges of M&E in an INGO. The projects I worked on were structured to develop a deeper understanding of those missed opportunities in the M&E system which Save the Children already outlined. Through my research I hope to confirm and validate the M&E challenges identified by Save the Children; I want to ask, with my limited understanding of Save the Children and the broader SC global movement, is there anything I can recommend to help SC strengthen the use and management of its global M&E system?

I started my research by learning about the overall organization of Save the Children. The following section summarizes my project work with Save the Children and provides useful background information about the organization and then concentrates on describing the components of its M&E system before moving into the findings of my research.

My involvement with Save the Children

I began work with Save the Children as Intern for Agency Performance Measurement (APM): Dashboard and Reporting, supervised by the Office of the President in Westport, CT. See Appendix B for an organizational chart showing this internship position. For three months I observed how the mission and strategic plan act as guideposts for agency-wide performance measurement efforts. The role of APM is to guide agency-wide assessments to illustrate significant successes and challenges. The agency-wide dashboard

> is used to measure progress towards the agency's multi-year strategic plan in support of our accountability for results objective. It is used to communicate w/staff and board on performance against our intended impact – our results for children as well as our theory of change (enable us to tell our story beyond the project/ program and country levels i.e. globally). It is also used as an organizational learning and business process improvement tool. (InterAction Forum, 2012, p.1)

I worked on the following projects during this internship: reviewing key financial metrics; designing and illustrating the top-ten metrics of the agency-level dashboard; reviewing and re-designing the 2012 annual benchmarking report; collecting and formatting quarterly reports from the Senior Management Team (SMT); and formatting the strategic plans for 2013-2015. The Office of the President uses APM to facilitate discussions among the SMT around agency-wide results, in particular through quarterly business review meetings. The president and chief executive officer (CEO) is ultimately accountable for agency-wide performance. The central M&E unit is responsible for developing the design, monitoring and evaluation standards and guidance, and for compiling and analyzing data at an agency-level related to Total Reach and program results.

After completing the APM internship I began another three month internship in the central M&E unit in Save the Children's office in Washington, DC. See Appendix C for an organizational chart showing this internship position. This unit provides guidance and

coordination for M&E and knowledge management (KM) across all areas of programming. Here the conversation around results converged on the merits and challenges of impact evaluations, efforts to promote shared learning and the utilization of evaluation findings. One of the major challenges central M&E unit helps to address is collecting, understanding, and synthesizing relevant data, like program results, so that various stakeholders can comprehend and appreciate the individual and cumulative results of Save the Children.

I worked on the following projects during this internship: interviewing M&E Specialists and personnel linked to the M&E system; documenting all program evaluations from 2010-2012, documenting all planned activities related to program evaluation for 2013; reviewing M&E technical resources; creating a brief for an impact evaluation known as PC3; and researching state-of-the-art M&E topics to present to the central M&E unit.

Introduction to Save the Children

There are three entities discusses throughout this paper: Save the Children, the United States based INGO; Save the Children International (known as SCI); and the Save the Children global movement, which includes all Save the Children member organizations around the world and SCI. The US-based Save the Children implements programs in about 60 countries. Its head offices are in Washington, DC, and Westport, CT. Save the Children's 2011 annual budget was around $600 million dollars (Save the Children, n.d.c). It is important to note that Save the Children is going through a major transformation in the way it works. Thirty separate Save the Children organizations around the world have created a new organization called Save the Children International (SCI) to serve as the program operations and delivery arm for their international programming. SCI requires standardization across organizations and will "create stronger global processes—including quality standards in areas like policy development,

program delivery, grant and financial management and accountability mechanisms—that maximize our collective capacity" (Save the Children, n.d.a, p.11).

A brief description of the key structures of Save the Children provides a frame for understanding the way this organization operates. The US-based Save the Children achieves results through six business teams which integrate staff from programming and functional departments. The business teams operate in the following sectors: Health and Nutrition (DHN), Education and Childhood Development (DECD), Hunger and Livelihoods (DHL), Humanitarian Response (DHR), Child Protection and HIV/AIDS (CP/HIV) and US Programs (USP). Representatives from functional divisions, like Public Policy and Advocacy, Finance, Fundraising, Human Resources, Information Technology and Legal Services, spread their technical expertise across the agency through the business team model to optimize and ensure the sustainability of the organization.

Save the Children's M&E system is a decentralized system. Although the central M&E unit coordinates and guides Save the Children's M&E policies and procedures, the practice M&E is not operated through the home office; country office personnel carry out program monitoring and conduct evaluations. The central Monitoring and Evaluation and Knowledge Management unit, with support from M&E and KM Specialists, provides leadership and technical support for the M&E and KM processes for the entire organization, especially country-based offices. Both of these responsibilities are known as cross-cutting functions because they encompass all programming areas. The central M&E unit also coordinates a Monitoring, Evaluation and Learning Working Group (MELWG) that advises international and US programs on agency-wide M&E issues and on practical ways to further enhance the quality, effectiveness and accountability of Save the Children's programs. The MELWG also guides Save the Children on

agency policy regarding cross-cutting programmatic, operational and capacity building M&E issues and takes action on those issues as appropriate.

Literature Review

The purpose of this literature review is to demonstrate how program evaluation grew into the specialized field of development evaluation and highlight the current state of its practice. The next part of this section will review the driving force for the effective use of development aid and some of the main challenges facing NGO's as they attempt to systematically monitor and evaluate their programs. The last piece of this section will frame the analysis of Save the Children's M&E system within its own standards for M&E and relevant standards from the Organisation for Economic Co-operation and Development (OECD).

The Growth of Development Evaluation

The practice of development evaluation is thought to have grown out of expansions in educational program evaluation stirring in public schools and universities. According to a study by Madaus, Stufflebeam and Kellaghan (2000) there have been seven stages in the development of program evaluation beginning in 1792 with the "first documented formal use of evaluation…[which] utilized the quantitative mark to assess students' performance (Hoskins, 1968)" (Hogan, 2007, p.3). Of course, master craft workers were judging their protégés handiwork and learned scholars were arbitrating arguments long before, probably thousands of years before, this stage described by Madaus et al. One place where evaluation became more formalized and formulaic was in manufacturing. The codification of evaluation took place step-by-step in response to rising demands for uniformity, replicability and accountability in the modernizing industrial world largely due to the US government vis-à-vis US military spending

(Hogan, 2007, p.4). The demand for military equipment for World War II led to improvements in the practice of evaluation just as development aid grew during the post-World War II era when the United States used it to help rebuild European states devastated by the war.

Education reform and education spending in the same post-World War II era brought increased attention to criterion-referenced testing and objective-based testing (OBT). OBT experienced a boom as "Bloom, Engelhart, Furst, Hill, and Krathwohl (1956) gave objective-based testing advancement when they published the *Taxonomy of Educational Objectives*" (Hogan, 2007, p.5). Furthermore, evaluation in education grew exponentially when "Congress enacted the [1958] National Defense Education Act (NDEA) which poured millions of dollars into new curriculum development projects… (Stufflebeam, Madaus, & Kellaghan, 2000). An additional push for program evaluation in education came with the 1965 Elementary and Secondary Education Act (ESEA) which "mandated the government to assess student performance and teacher quality standards and provided resources (first US government budgetary set-aside for evaluation) to undertake these activities, thereby institutionalizing evaluation" (IPDET Handbook Module 1, 2007, p.29). These congressional acts and the accompanying funds provided a much needed boost for the field of program evaluation.

The practice of development evaluation advanced in this post-World War II period as donor agencies began to recognize evaluation as an essential management tool and "began to emerge both at country level and in the UN system in the early 1950s. Since then, it has evolved slowly and unevenly" (UN Taskforce, 1984, p.7). The latter part of the 1950's and throughout the 1960's was a slow period of country level focus on M&E where the United Nations "promoted building of national development planning capabilities" (UN Taskforce, 1984, p.9). Building capacity in M&E was envisioned to increase ownership over the development process

for the governments and citizens in the countries where development programs were being implemented.

Interest and activities in developing M&E varied considerably during the 1960s and 1970s. Program monitoring and evaluation "such as in connection with development projects or technical assistance activities – were limited in concept and scope. They were concerned more with disbursement and delivery of physical inputs and outputs than with the nature of and impact on beneficiaries. In recent years, however, this has been changing (UN Taskforce, 1984, p.7). In the 1970's and 1980's, development evaluation "became a full-fledged profession in many OECD countries" and many monitoring and evaluation approaches, methods and standards were created (IPDET Handbook Module 1, 2007, p.30). It wasn't until the 1990's that development evaluation 'expanded and integrated' into professional associations with standards for the evaluation of international programs (Hogan, 2007, p.6). Now there are professional evaluation organizations, such as the American Evaluation Association, evaluation institutes, such as The Evaluators' Institute at George Washington University, and many online e-learning courses.

Development Evaluation in Practice

Monitoring and evaluation are defined differently based on an organization's needs. Evaluations are principally systematic assessments usually looking at program design, implementation or results set against a range of criteria. Save the Children and many others rely on the OECD to develop internationally recognized standards for the practice of monitoring and evaluation. Common types of development evaluations today are: *prospective* (scenario mapping for program design, forecasting costs and potential results), *formative* (the progress of a project being implemented), *summative* (final report on outcomes achieved), and *impact evaluation* (IE) (effect from the intervention continuing beyond the program end date) (IPDET Handbook

Module 1, 2007, p.13-15). Summative evaluations are the most commonly performed evaluation because they show immediate outcomes or project results. Formative evaluations like baseline assessments and mid-point evaluations are usually only used for larger or longer-term programs. Impact evaluations have become a desirable standard among large funders, typically governments, foundations or corporations.

Approaches for impact assessments from different fields of research have been lacking cohesion in their methods, "In recent years, there have been several efforts to integrate social and environmental impact assessments into more coherent forms (see Barrow 1997); impact analysis, on the other hand, was essentially confined to an assessment of impact several years after a project had ended" (1999, p.18). Roche emphasizes the challenges of theoretical approaches, methods, and techniques being used to uncover the social effects of change. Assessments that integrate social and environmental factors capture more of the impact from given contexts. Roche describes one innovation in the organization of evaluation planning as the "next generation of planning in international development agencies" called Logical Framework Analysis (LFA) or LogFrame (1999, p.18). LFA's attempt to "set out a clear hierarchy of inputs, activities, and objectives and to relate these to assumptions made about the external environment. Today the logical framework and its variants are the most common planning framework used by bilateral and multilateral agencies" (Roche, 1999, p.18-19). With impact evaluations, evaluators need to move beyond the project's LFA to determine whether the inputs, outputs, and outcomes caused the long-term change that planned. Impact analysis also requires the evaluator to determine whether the change can be attributed back to the implementing organization or if other factors were involved.

Some evaluations are specially structured to analyze systems and trace their processes. The systems thinking approach is an appropriate strategy for evaluation since international development organizations work globally, across all sectors and across many professions, academic disciplines and industries. Cabrera, Colosi & Lobdell (2008) note that systems thinking is suitable for complex programs specifically because it "is interdisciplinary and may act as a bridge between the physical, natural, and social sciences" and that it "may also act as a bridge between academic, professional and lay communities" (Cabrera, p.300-301). This means that this approach may also generate the most understanding between the actors in these different sectors.

M&E systems have developed within organizations to manage the development evaluation process. Directions from the OECD DAC on the institutional structure for managing evaluations asserts, "organisational aspects must address three requirements: developing a policy and a set of guidelines for evaluation; ensuring impartiality and independence; linking evaluation findings to future activities… by having a central unit responsible for evaluation reporting directly to…the agency head" (OECD, 1991, p.6). The OECD DAC further stipulates that an M&E system needs to have "a critical mass of professional evaluation staff in order to have sufficient expertise in their various fields of activity and to ensure the credibility of the process" (OECD, 1991, p.7).

M&E systems need a strategic plan to account for and coordinate agency-wide evaluation objectives and activities. OECD DAC directions for evaluation programming dictate the need for one overarching plan with a timetable for prioritizing and evaluating "a more comprehensive scale and an aggregation of evaluation results" (OECD, 1991, p.8). Furthermore, "credibility of evaluation depends on the expertise and independence of the evaluators…" and "evaluation should report successes as well as failures" (OECD, 1991, p.7). Since there are usually many

people with a stake in evaluations, it is important for the evaluation process to be transparent. In that light, one key component of an overall evaluation plan is its assurance of systematic dissemination of evaluation reports and a synthesis of results through annual reports, briefs, meetings and workshops in various settings. One significant area not addressed by the OECD DAC guidance is records management regarding evaluation reports. Evaluation findings are the evidence needed to support programs and advocate for policy changes. Key questions should be answered in evaluations so that "lessons learned and the recommendations provide the link between the results of the evaluation and future policy and programme development" (OECD, 1991, p.10). Most of all, M&E systems should facilitate feedback from stakeholders and link it to future activities. This ensures that future policies are developed with the concerns of stakeholders in mind and this increases the ownership of stakeholders within the development process.

The Push for Effectiveness and Accountability in Development Aid

"Since the 1990s there has been a major shift in the delivery of aid assistance away from donor designed and managed projects…associated with the end of the Cold War, theoretical critiques of development from the right and left, globalization, increased importance of trade and private investment, aid fatigue among donors and structural adjustment" (Conlin, 2008, p.194). The push for aid effectiveness and accountability drives the current emphasis on M&E in the development field. "[S]upport for development aid depends on the public's belief in its effectiveness. The moral case for providing support rests upon its achieving its objectives" (Roche, 1999, p.3). The funders of development programs, governments, foundations and charities alike, have a development agenda and need the implementing organizations to show that objectives are being met and that the overall human condition is being impacted as a result of development aid in order to justify the expenditure of funds. Recent shifts in development

assistance emphasize "impact and the effect of development assistance…this shift has been to move from project-based logical frameworks into a results framework" (Conlin, 2008, p.195). The objectives of the Millennium Development Goals agreed upon in 2000 exemplify the move toward managing for results and underscore the need for country-based ownership of development strategies. In 2005, leaders from around the world gathered to express their resolve to improve program performance through the formulation of what become known as the Paris Declaration on Aid Effectiveness:

> We, Ministers of developed and developing countries responsible for promoting development and Heads of multilateral and bilateral development institutions, meeting in Paris on 2 March 2005, resolve to take far-reaching and monitorable actions to reform the ways we deliver and manage aid as we look ahead to the UN five-year review of the Millennium Declaration and the Millennium Development Goals (MDGs) later this year. As in Monterrey, we recognize that while the volumes of aid and other development resources must increase to achieve these goals, aid effectiveness must increase significantly as well to support partner country efforts to strengthen governance and improve development performance. (United Nations, 2005, p.1)

These Ministers clearly state the justification for development aid rests on the ability of organizations to show real progress. These forerunners of development partnership "codified already emerging principles for development partners…donors should *align* their activities behind the country's strategy. This implies that donors work towards *harmonizing* their different approaches" (Conlin, 2008, p.195). As discussed in the Paris Declaration on Aid Effectiveness, the effectiveness of the work done by implementing organizations needs to be monitored and evaluated beyond a single project and those results need to be aggregated and reported periodically "because demonstrating real progress at country level is critical, under the leadership of the partner country" to "assess, qualitatively as well as quantitatively, our mutual progress at country level in implementing agreed commitments on aid effectiveness" (United Nations, 2005, p.2). This is difficult due in part to the complexity of evaluation approaches and methods as well as skepticism from critics of the NGO sector.

Current Challenges for Development Evaluation Systems

The practice of development evaluation is still maturing and evolving. For over 50 years the OECD has been working with governments to "understand what drives economic, social and environmental change" and promote "policies that will improve the economic and social well-being of people around the world" (OECD, n.d., Our Mission). Nonetheless some evidence suggests that "most current attempts to promote institutional learning, impact assessment, and greater accountability in the NGO sector" are insufficient (Roche, 1999, p.3). The context for development evaluation is facing new challenges caused by shifts in the structure of development assistance. Accountability has taken a new direction, downward, valuing beneficiary empowerment, in addition to donors, governments and civil society. The push for participation of beneficiaries in development evaluations is increasingly political and requires involving many more people. Moreover evaluation planning now needs to consider "managing for results and the Millennium Development Goals; partnership principles; new aid modalities such as sector approaches and general budget support; [and] going beyond aid to policy coherence" (Conlin, 2008, p.195). Roche continues to identify factors whose synergies contribute to breakdowns in development organizations:

> These elements combine to produce a growing gap between the rhetoric of agencies and the reality of what they achieve.

> Figure 1.1: The vicious circle
> - Increased competition between NGOs
> - Increased pressure to show results and impact
> - Lack of professional norms and standards
> - Poor learning and accountability
> - Growing need for profile for fundraising and [advocacy] work
>
> (1999, p.2)

These factors directly affect development evaluation systems and cause problems for M&E oversight and actually using the findings and recommendations from evaluations. Further issues arise when organizations cannot meet program evaluation needs and standards.

Not only are evaluations technically complex and multifaceted, so are M&E systems. The system that manages program evaluation needs to be evaluated itself. A study of the evaluation systems of OECD members found "that most development agencies now have a clear evaluation policy providing a clear mandate for the main evaluation unit and protecting the independence of the evaluation function" (OECD DAC NDE, June 2010, p.11). The study notes upward trends for joint evaluations and evaluation with wider scopes (OECD DAC NDE, June 2010, p.11). Some challenges persist, including "human resource capacities, supporting lesson learning and ensuring the systematic take up of evaluation findings to improve development effectiveness" (OECD DAC NDE, June 2010, p.11). An organization struggling to manage its M&E system might be aware of its system challenges or it may not. This is why it is essential to feature feedback mechanisms in the evaluation process:

> These would include such measures as evaluation committees, seminars and workshops, automated systems, reporting and follow-up procedures. Informal means such as networking and internal communications would also allow for the dissemination of ideas and information. In order to be effective, the feedback process requires staff and budget resources as well as support by senior management and the other actors involved. (OECD, 1991, p.11)

Organizations interested in enhancing the accountability effectiveness of their M&E system should review their monitoring and evaluation policies and compare them to 1) OECD standards and 2) their practices. Discrepancies between these sources point to challenges that need to be addressed.

Framework for Analysis

Save the Children describes the challenges it recognizes in its M&E system in its draft concept note *Evaluation and Learning Fund* with a section titled, "The Evaluation Gap: Staff Capacity and Financial Resources" (Save the Children, 4/30/12, p.2). Gaps are disparities formed by inconsistencies between M&E policies and procedures and M&E practices. The noted variances include: evaluations being underfunded; too few central M&E unit personnel and technical department personnel; insufficient capacity to document best practices and success stories; prioritizing program expansion and fundraising over evaluation and reflection; projects closing without proper documentation of lessons learned; lack of rigorous evaluations bringing evidence to support innovation about how and why interventions are working (Save the Children, 4/30/12).

Save the Children is working to close these gaps in its M&E system. During my internships learned about Save the Children's M&E system and its challenges. I conducted research to build on this knowledge base and develop a deeper understanding of challenges in the following areas: past evaluations, future evaluation plans, M&E tools and resource organization and challenges in developing Signature Programs. Furthermore, my analysis considers executive oversight of the M&E system. My analysis compares my findings to OECD standards and suggests measures to correct noted discrepancies.

Methodology

I learned about agency performance measurement and M&E through two separate three month internships in Save the Children. During both of these opportunities I was given specific projects and tasks to accomplish which contributed to this paper. My research primarily reviewed

M&E system documents, catalogued 2010-2012 evaluation reports and 2013 evaluation plans, questioned key home office level M&E system stakeholders and mapped M&E resources and tools. I compiled my findings and presented them to the MELWG. This research is intended to deepen Save the Children's understanding of the challenges it previously identified in its M&E system, and to propose realistic solutions for Save the Children to address those challenges.

The main methods of data collection for this inquiry were participant observation and key M&E system stakeholder interviews, primary document analysis including a mapping exercise of Save the Children's internal website, SaveNet, and its central M&E system files, plus independent secondary research into current state-of-the art M&E practices. The interviews conducted involved 14 individuals in the home office level either fully or partly responsible for supervising or coordinating M&E in Save the Children. The interviews targeted members of Save the Children's Monitoring, Evaluation and Learning Working Group (MELWG). See Appendix D for a list containing the positions of personnel interviewed. These interviews were conducted in support of the projects assigned to me during my internship. My supervisor and I developed the following key areas for the interviews:

I. Identifying and gathering all completed evaluations from your thematic area, 2010 to present
II. Reviewing the 2013 evaluation pipeline for your thematic area
III. Compiling a brief list of monitoring gaps and/or evaluation gaps that you can identify around signature program development in your thematic area
IV. Reviewing M&E tools and resources currently available on SaveNet for your thematic area

(Save the Children, December 2012, p.2)

To avoid redundancies or confusion, supplementary details about inquiry methods are given at the beginning of each section of analysis.

Limitations of My Exploration of Save the Children's M&E System

The scope of this research is intentionally limited to specific areas of concern in Save the Children's M&E system. Therefore there are several significant limitations to my analysis of Save the Children's M&E system. First, this analysis does not critique or discuss in detail the many approaches, methods and techniques of evaluation which Save the Children employs. Instead, the opinions of expert informants based on their experience in Save the Children are used to make generalizations about the focus areas in my research. Additionally, my analysis only incorporates evidence gathered at the head office level. The document review includes an analysis of outlined CO level evaluation plans. The scope of the document review did not include actual budgets for M&E within Save the Children. Budget information is limited to M&E policy guidelines for minimum funding standards for program evaluation.

Furthermore, my methodology did not include a direct systematic assessment of the skill level, knowledge base or experience of M&E personnel at head office or CO level. Consequently, there is no significant evidence from which to make a generalization about the internal capacity (skill level, knowledge base or experience) of M&E personnel at head office or CO level. Lastly, the final section of my analysis on M&E oversight is meant as an overview. This topic should be considered for stand-alone research since it connects to management theory as much as M&E.

Several key issues are not discussed due to the limited scope of this paper. For example, timeliness and the accessibility of evaluation findings by beneficiaries, beneficiary understanding and impartiality of evaluation plans which shape the evaluation's scope, questions and, of course, their findings. This paper does not trace how all evaluation findings are used nor does it consider all the possible consumers of evaluation reports. Those decisions are made by numerous

individuals within the organization and by donor stipulations, both of which are beyond the scope of this paper.

Analysis of Save the Children's M&E System

The following analysis focuses on Save the Children's organization-wide efforts at systematic program monitoring and evaluation, including its M&E policies and procedures and personnel structure for monitoring and evaluation.

Save the Children's M&E System Policies and Personnel Structure

The following section describes Save the Children's M&E system structures and the ways for coordinating M&E activities and functions including its purpose, policies and procedures, components and personnel. Save the Children's M&E policy sets standard operating procedures and minimum operating standards for program monitoring and evaluation. There are three components to the M&E system: 1. *Structure* (clearly defined roles and responsibilities at each level: country level, regional level, home office level and global initiatives); 2. *Procedures* (process steps and minimum standards); and 3. *Data* (indicators, Total Reach, qualitative information and evaluations) (Save the Children, 2012). The SOP defines the lines of accountability for each level and where additional resources can be found. Save the Children's Evaluation Handbook is "a step-by-step, practical guide to the evaluation process… to help managers, teams and technical specialists to ensure that evaluations are systematically designed and implemented to a high standard" (SCI, 2012, p.5). Policies in the MOS, "all innovative projects/programmes and those of strategic importance should be evaluated at the end of the intervention period, regardless of the budget. In declared emergencies, a real-time evaluation and Evaluation of Humanitarian Action must be conducted in accordance with the Emergency Rules

and Principles" (SCI, 2012, p.8). The handbook includes the policy on when to evaluate a program, chiefly based on the length of the program timeline and its budget, evaluation planning and design, as well as, how to implement and use the results of an evaluation. The evaluation thresholds show when a mid-term evaluation is needed as well as when external evaluators should be employed.

M&E policy states that learning from evaluations is expected to inform decisions and policy making within Save the Children, as well as its partners and other stakeholders. Save the Children defines evaluation as "the systematic and objective assessment of an ongoing or completed project, programme or policy and its design, implementation and results. The aim is to determine the relevance and fulfillment of objectives, development efficiency, effectiveness, impact and sustainability" (SCI, 2012, p.6). The Evaluation Handbook also makes clear that not all routine reviews that systematically analyse data qualify as an evaluation; a full program evaluation is different. The specific approaches (Planned, Results-based, Participatory), methodologies (case studies, ethnographic, participatory), techniques (interviews, surveys) and indicators (regarding inputs, outputs, outcomes, impact and sustainability) used in an evaluation plan are the responsibility of CO level staff (Save the Children, 2012, p.4).

The goals of Save the Children's M&E system are ambitious. The specific purpose for conducting program monitoring and evaluation is to "clearly articulate, demonstrate and document the outcomes of our work for girls and boys and their [caregivers]" (SCI, 4/2/12, p.4). Program monitoring data and evaluation findings contribute to institutional learning and accountability when they:

- Identify the outcomes and impact[3] of our work (at project, programme, country, regional or global level)
- Determine what helps and/or hinders our efforts to improve the quality and effectiveness of our programmes and projects
- Identify value for money and ensure that funds are used efficiently and effectively

- Strengthen a learning culture through reinforcing the links between evaluation findings and programme design and decision-making
- Validate our contribution to the key components of the theory of change
- Ensure accountability and transparency to our stakeholders, including children and their [caregivers]
- Provide robust evidence for policy and decision-making and for our advocacy, fundraising and external communications about the impact and value of our work
- Improve future project design and management, not just by proving impact but also by improving practice and by developing and documenting replicable, innovative solutions to the problems facing children.

[3] Save the Children uses the following definition of impact: the totality of effects produced by an intervention, whether they be positive or negative, intended or unintended, direct or indirect, primary or secondary. Note: in emergency interventions, effects may be assessed in terms of outcomes rather than impact.

(SCI, 2012, p.4-5)

Evaluation plans are interconnected with the evaluation budget. The decision whether or not to plan a full evaluation for a program is explained in the MOS. A rubric showing various program lengths and budgets notes whether an evaluation is required, not required or recommended. Further stipulation clarifies,

You should check donor requirements too, as these may mean bringing the thresholds forward. See the Evaluation Standard Operating Procedures for examples of who might initiate and/or commission an evaluation. The entity that commissions the evaluation is responsible for oversight of decisions and approvals at all stages. (SCI, 2011, p.8)

The SOP designates about 3-4% of a program's budget for evaluation efforts. Therefore budgets for M&E personnel and evaluations fluctuate based on various factors, most notably the amount of total funding per program. According to the SOP all program activity is expected to have progress reports, annual reports and final reports depending on the length of the program and its overall budget. Monitoring and evaluation budgets typically include lines for internal technical assistance, ongoing monitoring for data collection, travel, training, translation services, data analysis, report writing and the dissemination and publication of findings. Factors that increase the costs for evaluations can include hiring external consultants and research teams, baseline or mid-term reviews, surveys, establishing unique data collection systems or hiring additional internal M&E personnel. These additional costs could push an evaluation budget up to 10-15%

of a program's total budget. Actual departmental and country office budgets for M&E were not reviewed.

Country Level M&E Personnel

Country office level personnel and partner organizations in the field are the direct implementers of programs. CO's select and design of the most appropriate framework for programs and are responsible for monitoring and evaluating programs for Save the Children The MOS & SOP indicate that selection and design of appropriate approaches, methods, indicators and techniques for each program's monitoring and evaluation is ultimately the role of personnel at each CO. CO's often hire external firms or consultants to implement evaluations and do not depend on program personnel. Stakeholder engagement is also the responsibility of the CO's.

M&E is ultimately the responsibility of the CO program manager. Although all CO's are advised to have a central M&E officer some do not. Some CO's have multiple locations, usually an office in the capital city and field offices that transition per program needs. The number of M&E personnel or personnel with M&E responsibilities varies by country office. Many individual programs within each country have their own M&E staff at the program level in field offices. The personnel in country offices are mostly nationals of that country. The principle reason for this arrangement is to build up the capacity and utilize the talents of the people in the country where organizations are implementing programs.

Head Office Level M&E Personnel

Head office M&E personnel include the Director of the M&E Unit and nine (9) full-time M&E Specialists. There are also research staff and other personnel who work on M&E in some capacity. M&E Specialists provide the primary managerial and organizational oversight of M&E

in US-based Save the Children. The head of the M&E unit is situated in Washington, DC, and has oversight not only for monitoring and evaluation but for the knowledge management functions as well. M&E Specialists are technical advisors for monitoring and evaluation that are imbedded in each business team and some global campaign teams. Other personnel have significant M&E responsibilities as part of their job duties. This is especially true for staff in cross-cutting areas.

Regional M&E personnel have helped to coordinate and support M&E activity over the years. Although regional personnel have been a critical mechanism within the M&E system over the years, organizational shifts have relocated those duties to SCI which has to do with Save the Children's transition to one global movement. My observations did not involve regional personnel and document analyses did not reveal much about their activities. They were noted as responsible for supporting CO's M&E functions, including organizing and prioritizing evaluation plans and reporting results.

Partnerships and Capacity Building

Internal M&E capacity is supported by internal and external trainings, working groups and regional meetings, usually conducted by the M&E Specialists. M&E Specialists communicate directly with CO's to provide technical assistance and resources. They travel to CO's to hold trainings and assist with evaluations when necessary. M&E specialists meet regularly (presently once every other month) in a Monitoring, Evaluation and Learning Working Group (MELWG) to coordinate their work.

Planning and Coordination

Data from projects and programs abound. The challenge is often to determine what is the most meaningful to monitor and to report that is related to an organization's reach and impact. One component of Save the Children global M&E system is an estimation of the total number of people that were in some way touched by Save the Children's programming in a given year. This standard organization-wide data collecting procedure is known as Total Reach. Save the Children defines Total Reach as "the number of men, women, boys and girls who are directly and indirectly benefiting from Save the Children or any of its implementing partners' efforts" (Save the Children, 2012). Measures are taken to remove duplication in the count since some people might have benefited from multiple Save the Children programs. This can be a time consuming and confusing process for some, but those who have taken the time to understand the defined Total Reach methodology and incorporate the reach count into their ongoing monitoring systems find it to be a helpful method of understanding the general scope of programming in a given country, region or globally. Data on program participants is helpful because it can

> enable us to assess any changes in equality and discrimination over time and to understand which children are being included in – or excluded from – the benefits of a programme or project. This provides vital information for managers to make adjustments to programme strategies or implementation. Data are often disaggregated into these key categories: sex, age, disability status, area of residence (rural or urban), ethnic group, migrant status, refugee status, single/double orphan, child-headed household. (SCI, 2011, p.23-24)

Another component of SC's global M&E system is a set of *Global Indicators* (GI). GI's have been developed across they key thematic areas identified by the SC global movement, such as education, child protection, health, etc. They are intended for use in all countries, with the ability to aggregate data at the country level and beyond. At the time of my internship there were 16 GI's for development programs and six for humanitarian response (emergency relief) situations. As this is a relatively new component of SC's global M&E system, the indicators are

still under discussion and not all CO's are currently able to report on these indicators. The GI's are designed to be relevant to the majority of SC's programs in each thematic area. Note that each individual program has its own indicators in addition to GI's. SOP guidance states that the quantity of program indicators should be kept to a meaningful minimum that can still secure the necessary information to monitor the progress of a program. These indicators can be used for program evaluations, although new indicators may be designed for evaluations as well.

One more component of SC's global M&E system is the Advocacy Measurement Tool. This tool captures data about policy and advocacy efforts connected to Save the Children's programs. I reviewed this tool but chose not include an analysis of it for this paper.

Findings from Head Office Personnel Interviews

Save the Children has implemented hundreds of development programs over the years to make lasting change for children around the world. By monitoring and evaluating these programs, Save the Children plans to learn what, how and why interventions work in order to expand, scale up and replicate successful approaches. Programs are principally monitored and evaluated by country office level personnel in dozens of field offices around the world and supported by M&E Specialists and a small central M&E unit at the head office level. The central M&E unit determines organization-wide M&E policies and standards.

Based on my review of relevant documents, Save the Children has a clear evaluation policy and M&E policies and standards appear to be based on internationally recognized standards. Save the Children's evaluation handbook has thorough descriptions of its M&E policies and procedures and M&E is intended for all programming areas. Although M&E procedures are clearly documented, there is concern that country level personnel cannot effectively apply the guidance. It was not part of the scope of my research to assess the skill level

of country level personnel, their understanding of M&E policy and procedures or the general quality of evaluations they produce.

Each of the following sections emphasizes one focus area. The sections begin with a detailed description of how I began my research in that focus area followed by a reporting of my findings and my analysis. The Conclusions and Recommendations section synthesizes the findings and suggests further areas for analysis and possible actions to close the gaps identified here. Several useful appendices illustrate the findings discussed below.

I. Focus Area: 2010-2012 Evaluation Reports

The completed evaluation report cataloguing project shows the M&E system's effectiveness at documenting its past evaluation efforts. This can be considered a proxy for the collective capacity for learning from past evaluations which is partly a knowledge management issue. I did not find a procedural requirement in OECD standards regarding the central storage of evaluation reports. The central M&E unit is interested in developing a central repository for all program evaluations which can become a building block for an evaluation library for the SC global movement. I learned the current expectation in the US-based Save the Children is that business teams store evaluations and other significant documents in document libraries on the internal website, SaveNet. Subsequently I began a catalogue of 2010 – 2012 reports by searching document libraries on SaveNet. Next I explored central M&E unit folders on the internal computer network. With a preliminary list of reports and evaluations assembled, I initiated interviews with M&E Specialists and other key staff. I sent the list for each interviewee to review prior to our meeting. During the interview I requested evaluation reports not on the list or final program reports for programs that did not have evaluation reports. In most cases, follow up

was needed to collect additional reports. By the end of my internship period, I collected fifty-four (54) evaluations and final reports. I catalogued these reports in a Microsoft Excel spreadsheet and analyzed the resulting list to find any patterns or trends in the data.

The catalogue of 2010 – 2012 reports includes categories for the year and country where they occurred, the initiating business team and sub-project, the title and the type of report. I sorted the reports by year which showed that half of the 54 reports collected are from 2011, while 33% occurred in 2012 and 17% occurred in 2010. The originating business teams for the majority of reports collected are DECD at 37% and DHN at 31%; a moderate amount came from CP/HIV at 13%; and the least came from DHL at 2%, DHR at 7% and USP at 9%. I divided the report types into nine categories. Evaluation Reports were 42% of all reports collected. Four report types are evaluative, including final evaluation at 19%, impact evaluation at 4%, mid-term evaluation at 13% and real-time evaluation at 6%, totaling 42%. The other report types are not evaluative, including assessment at 7%, baseline at 2%, final reports at 24%, results reports at 9% and study at 17%. See Appendix E for graphic representations of my findings.

According to my interviews and observations, evaluations are not systematically uploaded to SaveNet by most of the business teams. Also, there was a former effort to standardize evaluation report storage in a database called SaveIt. I was told that effort was not successful. Presently, many reports and evaluations are stored on individual computers or shared-drive folders on the internal network. Consequently some interviewees said they did not know how or where to locate evaluations held by their predecessors. After following up a few times over several weeks, they still could not locate evaluations from 2010 – 2012. It is possible that there were no evaluation reports from that period since there is a tendency for business teams rely mostly on final reports from program staff instead of initiating formal evaluations.

Interviews and documents reviewed confirm that the storage and organization of evaluation reports does not meet organizational expectations or needs. There is no central repository of evaluation reports although the central M&E unit is interested in archiving evaluation reports. Save the Children has established a culture of knowledge management throughout the agency and the SC global movement which acknowledges this issue and is in the process of addressing it.

II. Focus Area: 2013 Evaluation Plans

The future evaluation plan cataloguing project shows the M&E system's effectiveness at documenting its future evaluation efforts. This can be considered a proxy for the collective capacity for generating and gathering evidence from future evaluations. I did not find any guidance from the OECD for central documentation of evaluation plans. M&E plans for the upcoming year are documented in the CO's annual planning process and uploaded around September to SCI's internal website, XtraNet. CO's are now managed by SCI. Program evaluation plans are not centrally controlled or coordinated by the US-based Save the Children home office; its central M&E unit is in the process of developing a central pipeline where all program evaluation plans can be collected in one document which can become a building block for a global evaluation pipeline for the SC global movement.

I began to catalogue of all 2013 evaluation-related activities with the evaluation plans outlined in each country office's annual plan document. I separated out evaluations funded by Save the Children US to narrow the list before reviewing it with the M&E Specialists. Next I asked each interviewee to review the list and add any further evaluation plans they knew about. This effort generated 36 additional evaluation-related activities planned yet not included in country annual plans. The final document contains 125 planned evaluation-related activities. I

catalogued these evaluation plans in a Microsoft Excel spreadsheet and analyzed the resulting list to find any patterns in the data.

The catalogue of 2013 evaluation plans includes categories for the country where they will occur, the initiating business team and sub-project, the title, the type of evaluation activity, the source of funds and other notes. I sorted the plans by business teams and recorded the number and percent of evaluation activities being initiated by each team across Save the Children US and throughout the entire organization. The most evaluations were planned by DECD 65 (52%); a moderate amount came from DHN at 27 (22%) and DHL at 20 (16%); and the least came from DHR at 7 (6%), CP/HIV at 2 (2%) and Other/Not Labeled at 4 (3%). See Appendix E for graphic representations of these findings. The graph contains additional information, comparing the percent of all evaluation plans to the US-based Save the Children's evaluation plans in each thematic area.

US Programs (USP) did not provide their evaluation plans for 2013 so that business team's plans are not included in my graphs. Moreover my interviewee did not know of a document with aggregate evaluation plans for this business team. Domestic programs operated by USP do not always operate within the global M&E system. However, as the transition to one SC global movement continues, this topic will be given further consideration.

Primary document review showed that almost all country offices wrote something about their evaluation plans as part of the country annual planning process; thus evaluation plans were documented as expected. However, the level of detail in these evaluation plan summaries is skeletal and often incomplete. The US-based Save the Children needs to access the evaluations from nearly 60 CO's. In total, to be used by SCI, approximately 120 separate evaluation plan documents (one from each SCI member country office) need to be downloaded and then

combined before they can be analyzed in one document. This evidence indicates that the central mechanism that captures planned M&E activities is inefficient.

III. Focus Area: Gaps in Evidence Around Signature Programs

One of Save the Children's strategic goals by 2015 is to "take at least six Signature Program[s] to scale through policy change and partnerships. We will invest in evaluation and refine our approaches with what we learn, working with others to bring about high-quality results for children" (Save the Children, n.d.a, p.2). Programs advance through the Signature Program Pipeline after passing through three stages: Development, Pilot, Ready to Scale. There are fifteen programs in the pipeline according to this diagram. See Appendix F for a diagram of this pipeline that includes all the program names. The level of details with which the interviewees were able to describe their knowledge of M&E gaps in the evidence base of Signature Programs represents the level of coordination and communication toward these endeavors at the outset of the strategy period. This focus area was chosen since Save the Children was preparing to identify some program concepts to develop and build an evidence base in the first quarter of 2013.

I began my questioning of each informant by asking them to identify and discuss monitoring gaps and evaluation gaps around signature program development in their business team or thematic area. I took notes and logged their comments immediately following each interview. I made a qualitative review of my notes to discover themes among interviewee responses.

Discussions about Signature Programs quickly moved to include gaps in areas outside of Signature Program development. Most interviewees commenting on Signature Program development claimed to understand the gaps in evaluation evidence and to have an agenda to address them. Consequently there is a lack of substantive data for generalization. One

interviewee noted a need for mid-term evaluations in the course of programs with longer timelines. This concern was coupled with a need for more qualitative data to enhance the contextual evidence needed when scaling up programs. Another interviewee noted difficulty accessing raw program data and reports as well as questions about the quality of analysis of the reports. A different interviewee noted that it is challenging to negotiate with evaluations partners to ensure Save the Children's interests are met by evaluations. Research and evaluation partnerships are established over many years and many factors change, including priority learning objectives.

In addition to Signature Program development, interviewees discussed general challenges in the M&E system and possible solutions. For example, some interviewees noted that M&E Specialists do not consistently receive copies of summative evaluations or final program reports. They said this happens for two reasons. First, there is no central tracking mechanism for evaluations. Second, either projects do not have an assigned M&E Specialist or they do not have an evaluation planned and budgeted.

Workload is another challenge noted by various interviewees. They specifically pointed out that there are too many administrative tasks and responsibilities for M&E Specialists. Moreover, interviewees felt that excessive turnover has contributed to a loss of institutional knowledge and limits opportunities for collaborative research and evaluation partnerships. I noted two M&E Specialists who recently left their positions plus a couple more areas that formerly had no M&E Specialist.

Most interviewees acknowledged a limited or inadequate relationship with country office level personnel. Several interviewees emphasized that an incomplete understanding of the skill and knowledge level of country office level personnel makes it difficult for M&E Specialists to

provide technical assistance. Most interviewees feel that country office level personnel do not have the necessary capacity to conduct a rigorous program evaluation. One interviewee thinks that outsourcing evaluations to external consultants contributes to the evaluation capacity of country office level personnel remaining low. This interviewee also noted that evaluations are often written in English which is usually not the first language of country office personnel. All of these factors influence the level of ownership country office personnel have for evaluation results and their willingness and ability to communicate those results with local stakeholders.

The interviewees did not offer detailed descriptions of gaps in evaluation evidence around Signature Programs as requested. However they did indicate that they know about gaps and have plans to address those gaps. Some programs are designed to generate evidence toward Signature Programs and these internal interests form the basis and need for evaluation planning in those programs. Interviewees were able to describe general gaps in the M&E system in their programming areas but this was not the main focus of my investigation.

IV. Focus Area: M&E Tools and Resources

M&E Specialists provide assistance in many ways, including technical tools and resources. The level of organization demonstrated by the M&E resources and tools mapping project characterizes the M&E system's capability to provide efficient technical support. I did not find any guidance from the OECD on the organization of technical resources and tools. I began a review of M&E tools and resources by searching document libraries in the central M&E area and each business team area on the internal website, SaveNet. I mapped out the storage areas and illustrated their utility using PowerPoint. See Appendix G for a sample image from the resource mapping exercise. I reviewed my findings with each interviewee and asked about resource management and utilization by country offices.

The document libraries on SaveNet are supposed to be the central storage place where each business team maintains resources and tools for country office level staff as well as colleagues at the head office. The mapping exercise revealed that few business teams organize and update M&E tools and resources in their document libraries. Several document libraries were empty or out of date. One document library was up-to-date and well organized.

Interviewees said the M&E resources and tools they use the most are stored on their individual computers or in shared-drive folders on the internal network. There is a tendency for business teams to rely mostly on ad hoc arrangements to exchange resources. They are mostly exchanged via email on a case-by-case basis as requests for technical assistance are made by country office personnel or head office colleagues. Some interviewees noted trouble locating and using M&E resources and tools held by their predecessors. All together these findings indicate that resources for M&E are not properly catalogued. It is inefficient for dozens of country office level personnel with similar needs to make individual requests for technical support instead of accessing a well-organized database of resources.

The mapping exercise demonstrates that M&E technical resources and tools are not uniformly organized and the interviews confirmed this. Interviewees overwhelmingly felt that document libraries were not accessible and therefore not utilized by country office level staff. Interviewees noted that country office personnel lost access to SaveNet once they changed over management to SCI. Therefore the people primarily responsible for conducting M&E, country office personnel, did not have adequately to up-to-date and organized M&E resources. Further confounding this issue, the XtraNet, the internal website hosted by SCI, has few M&E resources and tools available since it is being constructed posthumously. The XtraNet has undergone further development since my internship and is now called OneNet.

Executive Oversight of M&E

At Save the Children, "senior leadership at all levels sees the value in using the dashboard to communicate better internally and externally about organizational progress against plan" (InterAction, 2012, p.3). The senior management team (SMT) is responsible for executive oversight at Save the Children. The SMT is made up of the top leaders of the organization, essentially the President and CEO, COO and the Vice Presidents. This team monitors major activity in business teams and functional divisions using dashboards updated every quarter. The BT dashboards roll up to form one agency-wide dashboard. The agency dashboard is made of the top ten agency-wide metrics followed by metrics based on the top objectives of business teams and functional divisions. The initial driving force of the agency dashboard is the "need to have a tool for management purposes from the senior leadership of the organization (including the board)" (InterAction, 2012, p.2). This dashboard does not contain a metric based on the use of evaluation findings; it does include the annual measure of Total Reach. The business team dashboards provide executive oversight of M&E by means of "total reach data and evaluation results in an annex" (Save the Children MELWG, 2/9/12, p.1). Reach is only a count of participants directly and indirectly affected by programs. Since these dashboards don't measure evaluation findings, they are a weak means of executive oversight. A full review of the agency dashboard suggests that executive oversight of the M&E system through the dashboard tool is limited. M&E system performance metrics should be part of the dashboard to provide more oversight of the M&E process.

At the time my internship in APM completed, Save the Children was in the process of developing a new agency-level dashboard which would incorporate metrics for its 2013-2015 strategic objectives. The business team dashboards that contribute to the agency dashboard were

set to be the home-base for more result indicators. My analysis does not consider the 2013 dashboards since they were not available for my review during my internship. The focus moving forward appears to be set on bringing program results to the forefront for wide sharing internally and externally and completing the transition to one SC global movement by harmonizing systems and building the capacity of CO's to use them.

Conclusion and Recommendations

Overall, the findings of my research confirm and validate the M&E system needs already identified by Save the Children. Although some evidence suggests additional weaknesses in the M&E system, I could not investigate them in the limited time of my internships in Save the Children. Taken as a whole, this research gave me valuable real-world exposure to the challenges of M&E in an INGO and affirms the need for more attention and research in this area. Methodical assessments that map disruptions in system processes and aim to recognize inefficiencies in feedback mechanisms can show the challenges to institutional capacity for M&E and related knowledge management. More often, these challenges exist despite decent people diligently working toward solutions under difficult constraints that stretch well beyond their organization's practice of M&E. This section summaries my findings and advocates on behalf of further analysis and recommended courses of action.

The primary document review and interviews demonstrated that there is not uniform storage of evaluation reports. Based on my findings evaluation plan summaries were catalogued as expected but those products were of limited use. Informants did not discuss details regarding gaps in evidence around Signature Programs. Alternatively, informants suggested that a significant amount of country level personnel did not have the skill level or necessary support to

adequately carry out M&E policies and procedures. It is unclear whether there was sufficient training for country level personnel. Access to technical resources and tools is one issue while applying them is a different issue altogether. The evidence showed poor access to and organization of the M&E technical resources and tools needed to plan and conduct program monitoring and evaluation. Following Save the Children's example plus utilizing the following recommendations, organizations can develop their M&E system by proactively address challenges.

Recommendations

These findings represent a review of Save the Children's M&E system focused on key areas of concern. Additional reviews might include systematically matching up past programs to with evaluations. This mapping activity should go back about 10-15 years in order to capture a significant amount of evaluations. Furthermore analyzing program evidence and grant applications to show how often evaluation findings are cited. It would also be useful to note whether those findings come from Save the Children evaluations or other organizations.

Save the Children could expand the functionality of its central M&E unit as a means to achieve more checks-and-balances and accountability within its M&E system. This expanded functionality could increase executive oversight of M&E. Additional functions could include formalizing and enhancing the process for central evaluation report storage, evaluation product development, technical resource storage, plus rating various elements within the M&E system. In order to develop these extra functions, additional investment would be needed. Currently the funds for the strategic investment are reserved for evidence generation, documentation and sharing, as approved by the SMT and the agency's board. Future investments could be focused on developing metrics and increasing oversight of the M&E system.

Developing metrics to rate the M&E system and using them on dashboards will increase accountability within the system, as well as executive oversight. Those metrics could also become part of an external benchmarking and peer review process. That process could include an audit of its M&E system using a peer-review assessment method developed by the OECD and published in *Evaluating Development Co-operation: Summary of Key Norms and Standards* (OECD DAC NDE, June 2010). The OECD DAC framework for assessing development evaluation systems can be tailored to the unique needs to any peer set.

> This framework was…designed to strengthen the evaluation function and promote transparency and accountability in development agencies. It has been developed with peer reviews in mind and as a management device for improving evaluation practice in aid agencies. It is a "living" tool, meant to be updated in function of experience. (OECD DAC NDE, June 2010, p.15)

This framework can be used to audit M&E systems on a regular basis. It can be used internally or as a peer review mechanism which could allow for agency-level benchmarking. This would be a way for Save the Children to lead and innovate in the field of M&E.

One measure of the use of evaluation findings could be evidence from evaluation reports cited for new programs. Another measure might be timeliness of evaluation reports and feedback. The central M&E unit could also commission a meta-analysis of evaluation reports to objectively determine their overall quality. Additional ideas for areas to measure can be found in *Evaluation Feedback for Effective Learning and Accountability* (OECD, 2011). Forming metrics requires technical knowledge and an intricate understanding of policies and procedures currently in use. M&E Specialists and the MELWG could be engaged to champion the development of new metrics. Establishing a portfolio of M&E related metrics will not only enhance executive oversight of M&E, it will increase effectiveness of the M&E system and hopefully expedite the process of creating breakthroughs in the way the world treats children and achieve immediate and lasting change in the lives of boys and girls the world over.

References

Cabrera, D., Colosi, L., & Lobdell, C. (2008). Systems thinking. *Evaluation and Program Planning 31*, 299-310.

Conlin, S. and Stirrat, R. (2008). Current challenges in development evaluation. (pp.193-208) *Evaluation 14(2)*. Sage Publications.

Hogan, R. L. (2007). The historical development of program evaluation: Exploring the past and present. *Online Journal of Workforce Education and Development 2(4)*. Eastern Illinois University.

Independent Evaluation Group and Carleton University. (2007). Module 1: Introduction to development evaluation. (pp.9-30) *International Program for Development Evaluation Training (IPDET) Handbook,* The World Bank Group.

OECD. (n.d.) About the OECD. (Online) Retrieved from: http://www.oecd.org/about/ Accessed January 14, 2012.

OECD DAC. (1991). Principles for evaluation of development assistance. Paris, France. Retrieved from: http://www.oecd.org/dac/evaluation/50584880.pdf. Accessed January 14, 2013.

OECD DAC Network on Development Evaluation [OECD DAC NDE]. (June 2010). Evaluating development co-operation: Summary of key norms and standards. (2nd ed.). Paris, France. Retrieved from: http://www.oecd.org/dac/evaluation/dcdndep/41612905.pdf. Accessed January 14, 2013.

Roche, C. (1999). Impact assessment for development agencies: Learning to value change. Oxford, UK: Oxfam GB.

Save the Children. (n.d.a). Save the Children 2013-2015 strategic plan: Investing in impact for children. Retrieved from: http://www.savethechildrenweb.org/Investing-in-Impact-for-Children/. Accessed December 20, 2012.

Save the Children. (n.d.b). About us. (Online). Retrieved from: http://www.savethechildren.org/site/c.8rKLIXMGIpI4E/b.6146405/k.C7E9/About_Us.htm. Accessed December 20, 2012.

Save the Children. (n.d.c). Save the Children 2011 IRS form 990. Accessed December 20, 2012. Retrieved from http://www.savethechildren.org/atf/cf/%7B9def2ebe-10ae-432c-9bd0-df91d2eba74a%7D/STC_990_2011.PDF

Save the Children. (n.d.d). Who we are. (Online). Retrieved from: http://www.savethechildren.org/site/c.8rKLIXMGIpI4E/b.6146355/ Accessed December 20, 2012.

Save the Children. (April 2012). Evaluation and learning fund: Draft concept note. Word document. Accessed October 25, 2012.

Save the Children. (May 2012). Results for children Q1-2 2012. Retrieved from: http://www.savethechildren.org/atf/cf/%7B9DEF2EBE-10AE-432C-9BD0-DF91D2EBA74A%7D/Results-for-Children-Q1Q2-2012.pdf. Accessed December 2, 2012.

Save the Children. (5/27/10). Save the Children statement of approach to monitoring and evaluation. Word document. Accessed November 5, 2012.

Save the Children. (10/22/12). SaveNet M&E Mapping- Scott. PowerPoint document. Accessed November 12, 2012.

Save the Children. (12/13/12). Interview Log. Word document. Accessed December 17, 2012.

Save the Children. (12/14/12). Interview Findings Report. PowerPoint document. Accessed
December 17, 2012.

Save the Children International [referenced as SCI]. (4/2/12). Save the Children evaluation
handbook. PDF document. Accessed December 2, 2012.

United Nations ACC Task Force on Rural Development, Panel on Monitoring and Evaluation
[referenced as UN Taskforce]. (1984). Guiding principles for the design and use of
monitoring and evaluation in rural development projects and programmes. Rome. (ACC-
Administrative Committee on Coordination)

United Nations. (2005). Paris Declaration for Aid Effectiveness. Retrieve from:
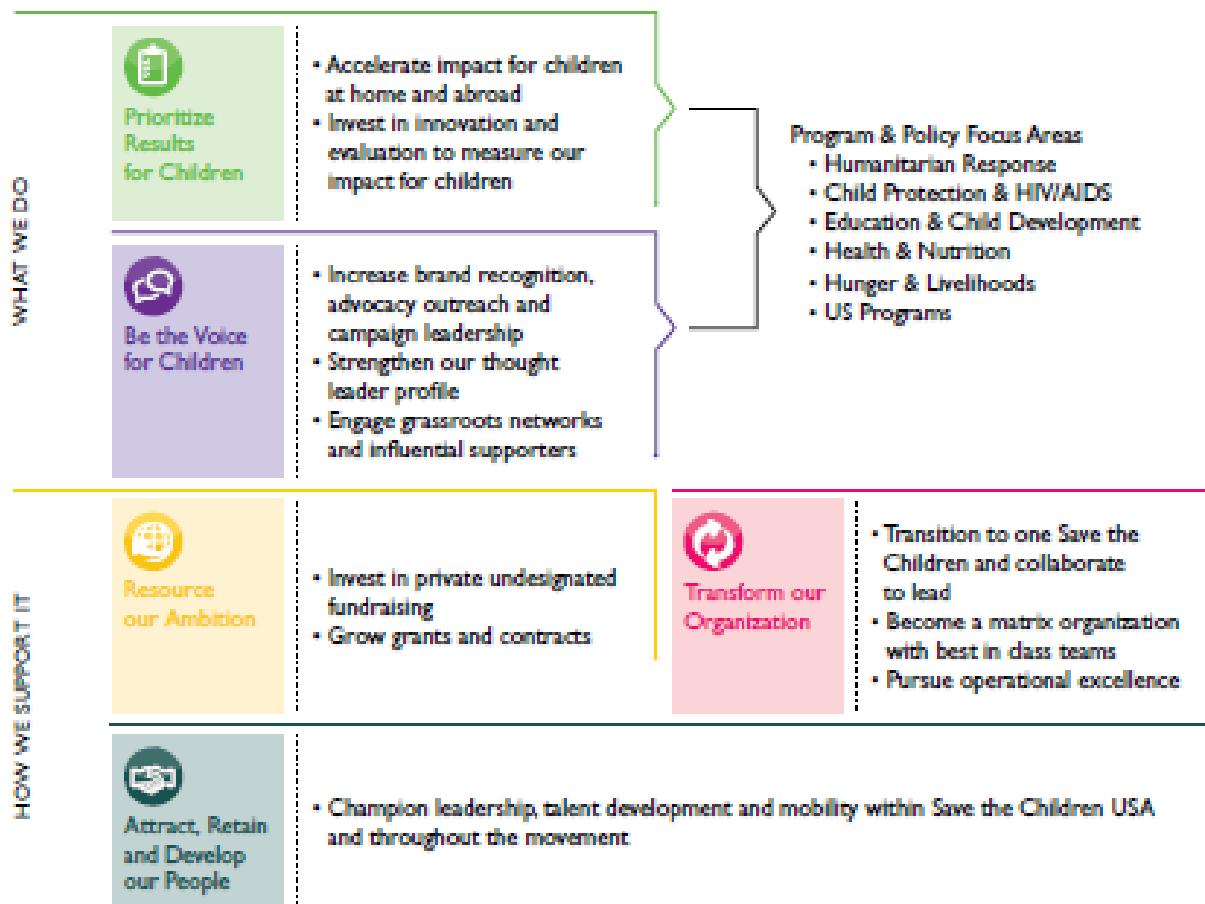http://www.oecd.org/development/evaluation/dcdndep/38245246.pdf

Appendix A

Save the Children 2013-2015 Strategic Plan: Investing in Impact for Children (Save the Children, n.d.a, p.1)

Investing in Impact for Children

Save the Children 2013-2015 Strategic Framework

By 2015, we will have successfully transitioned into one global movement, driven our Theory of Change in more countries and contributed significantly to achieving breakthroughs for children.
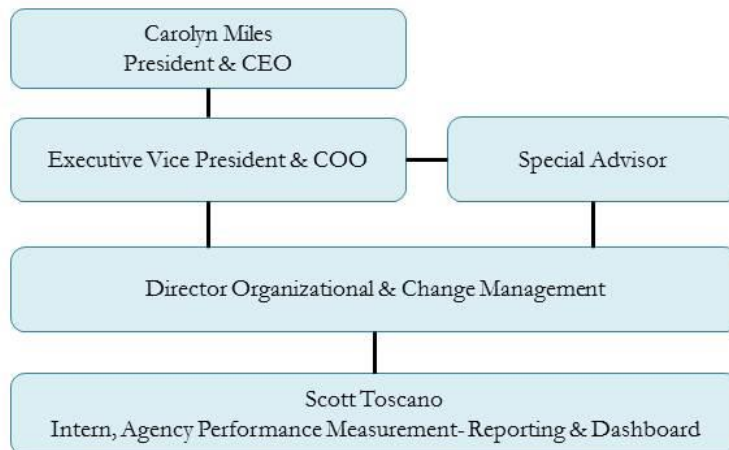
WHAT WE DO

Prioritize Results for Children
- Accelerate impact for children at home and abroad
- Invest in innovation and evaluation to measure our impact for children

Be the Voice for Children
- Increase brand recognition, advocacy outreach and campaign leadership
- Strengthen our thought leader profile
- Engage grassroots networks and influential supporters

Program & Policy Focus Areas
- Humanitarian Response
- Child Protection & HIV/AIDS
- Education & Child Development
- Health & Nutrition
- Hunger & Livelihoods
- US Programs

HOW WE SUPPORT IT

Resource our Ambition
- Invest in private undesignated fundraising
- Grow grants and contracts

Transform our Organization
- Transition to one Save the Children and collaborate to lead
- Become a matrix organization with best in class teams
- Pursue operational excellence

Attract, Retain and Develop our People
- Champion leadership, talent development and mobility within Save the Children USA and throughout the movement

Appendix B

Organization Chart Showing Position: Intern for Agency Performance Measurement

Appendix C

Organization Chart Showing Position: Intern for Monitoring and Evaluation
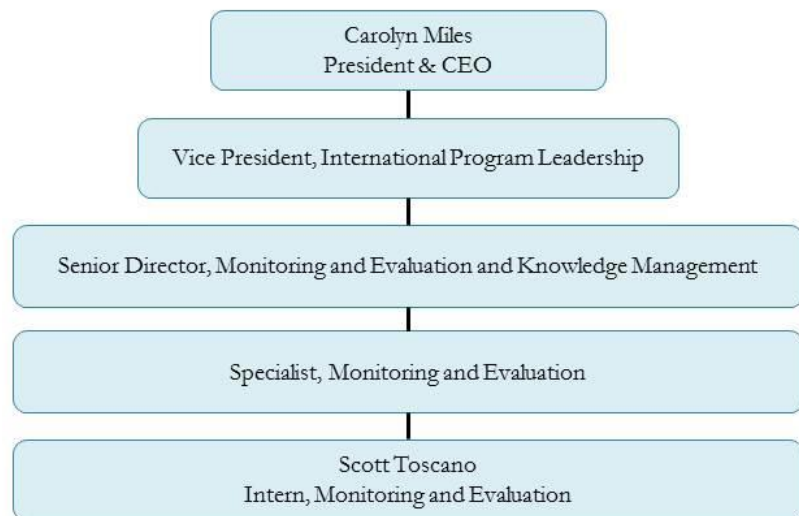
Appendix D

<u>List of Personnel Interviewed and Interview Focus Areas</u>
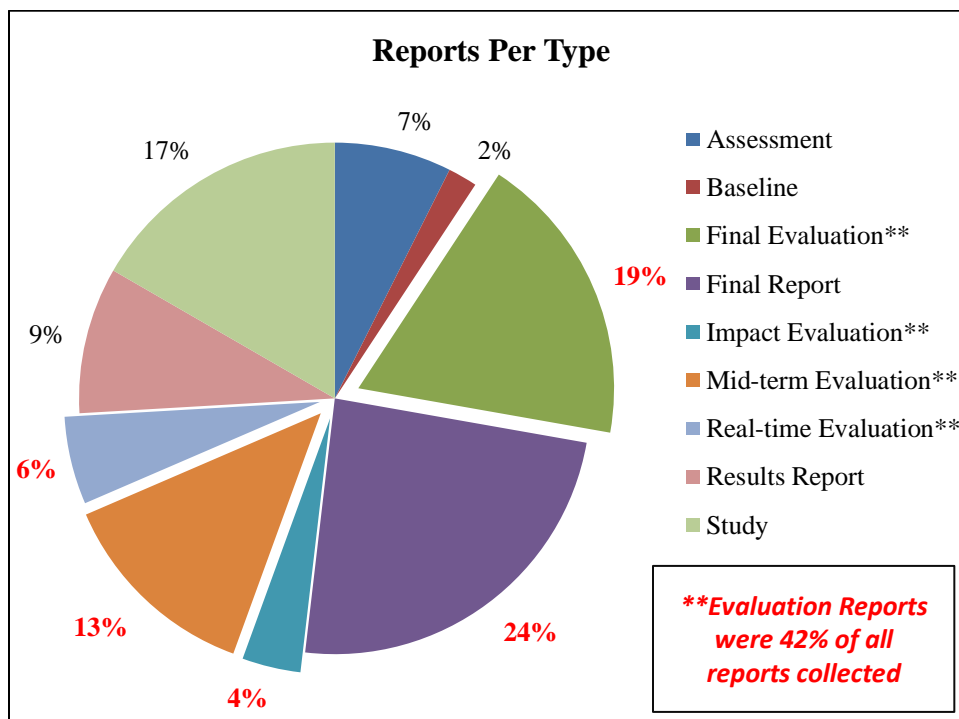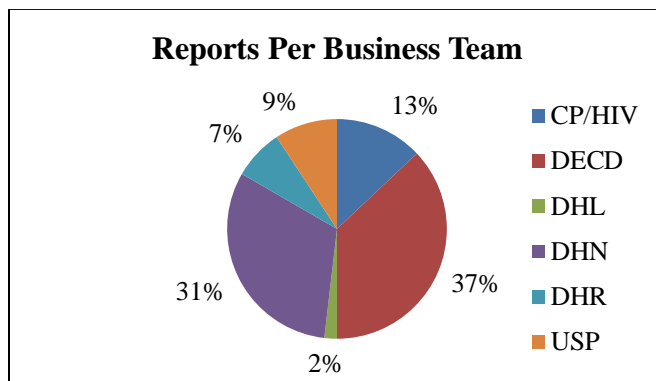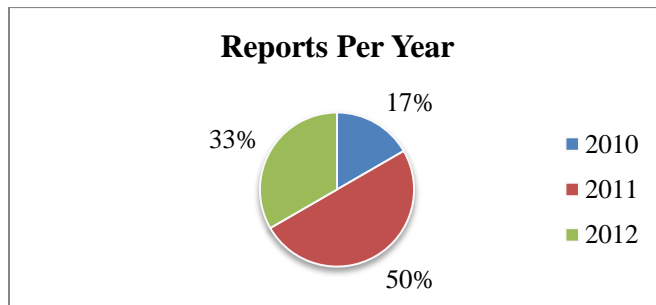
Interview Focus Areas
1. Identifying and gathering all completed evaluations from your thematic area, 2010 to present
2. Reviewing the 2013 evaluation pipeline for your thematic area
3. Compiling a brief list of monitoring gaps and/or evaluation gaps that you can identify around signature program development in your thematic area
4. Reviewing M&E tools and resources currently available on SaveNet for your thematic area

List Personnel Interviewed: November to December 2012

| Position	Department
Associate Director DM&E and Research	DHL
Coordinator, Portfolio Documentation, Communication & Support	DHL
Sr. M&E Specialist	CP & HIV
M&E Specialist	DHR
M&E Specialist	DHN, SNL
Senior M&E Specialist	DHN
M&E Specialist	DECD- Literacy Boost
Senior Director	DECD
Specialist Information & Documentation	DECD
Director	GIRL Project
Director, Field Policy and Advocacy	PPA
Data Analysis & Reporting Specialist	USP
M&E Specialist	DHN, SNL Campaign
M&E Specialist	DHN, Emergency Health and Nutrition
Advisor	Sponsorship Funded Programs

Appendix E

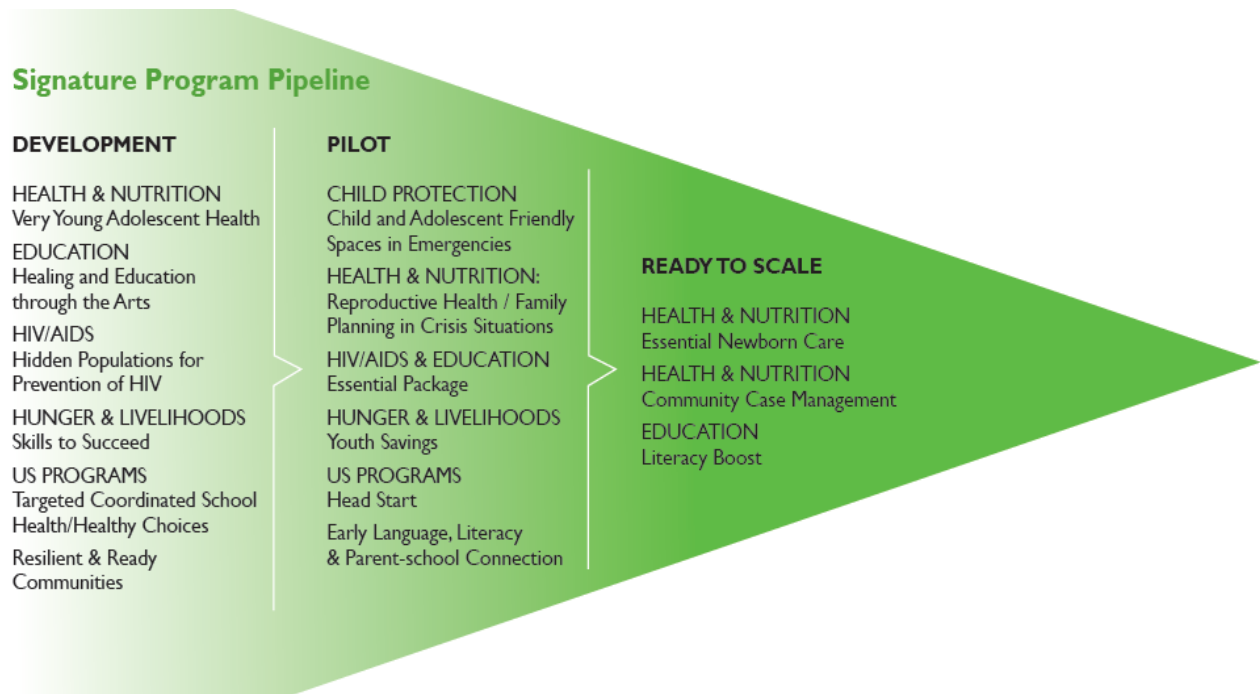Graphs of 2010 – 2012 Reports Catalogued (Save the Children, 12/14/12, p.1)

Reports Per Year

Reports Per Business Team

Reports Per Type

**Evaluation Reports were 42% of all reports collected

Appendix E continued

Graph of Planned Evaluations for 2013 Catalogued (Save the Children, 12/14/12, p.2)

Percent of Evaluations by Theme in 2013 Pipeline
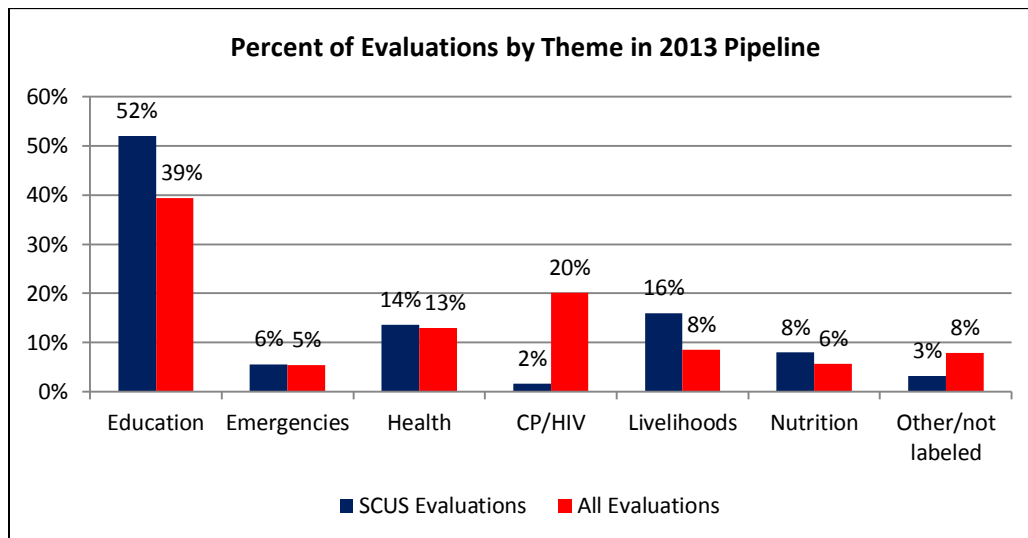
Appendix F

Save the Children's Signature Program Pipeline (Save the Children, n.d.a, p.2)

Signature Program Pipeline

DEVELOPMENT

HEALTH & NUTRITION
Very Young Adolescent Health

EDUCATION
Healing and Education
through the Arts

HIV/AIDS
Hidden Populations for
Prevention of HIV

HUNGER & LIVELIHOODS
Skills to Succeed

US PROGRAMS
Targeted Coordinated School
Health/Healthy Choices

Resilient & Ready
Communities

PILOT

CHILD PROTECTION
Child and Adolescent Friendly
Spaces in Emergencies

HEALTH & NUTRITION:
Reproductive Health / Family
Planning in Crisis Situations

HIV/AIDS & EDUCATION
Essential Package

HUNGER & LIVELIHOODS
Youth Savings

US PROGRAMS
Head Start

Early Language, Literacy
& Parent-school Connection

READY TO SCALE

HEALTH & NUTRITION
Essential Newborn Care

HEALTH & NUTRITION
Community Case Management

EDUCATION
Literacy Boost

Appendix G

Sample of M&E Resource Mapping Exercise from file "SaveNet M&E Mapping- Scott" (Save the Children, 10/22/12, slide 6)

Appendix H

Part III- Evaluation Systems and Use: A Working Tool for Peer Reviews and Assessments* from Evaluating Development Co-operation: Summary of Key Norms and Standards (OECD DAC NDE, June 2010, p.15-18) *Formatting has been altered, chiefly the use of numbering instead of bullet points.*

I. Evaluation policy: role, responsibility and objectives of the evaluation unit
1. Does the ministry/aid agency have an evaluation policy?
2. Does the policy describe the role, governance structure and position of the evaluation unit within the institutional aid structure?
3. Does the evaluation function provide a useful coverage of the whole development cooperation programme?
4. According to the policy, how does evaluation contribute to institutional learning and accountability?
5. How is the relationship between evaluation and audit conceptualised within the agency?
6. In countries with two or more aid agencies, how are the roles of the respective evaluation units defined and coordinated?
→ *Is the evaluation policy adequately known and implemented within the aid agency?*

II. Impartiality, transparency and independence
7. To what extent are the evaluation unit and the evaluation process independent from line management?
8. What are the formal and actual drivers ensuring/constraining the evaluation unit's independence?
9. What is the evaluation unit's experience in exposing success and failures of aid programmes and their implementation?
10. Is the evaluation process transparent enough to ensure its credibility and legitimacy? Are evaluation findings consistently made public?
11. How is the balance between independence and the need for interaction with line management dealt with by the system?
→ *Are the evaluation process and reports perceived as impartial by non- evaluation actors within and outside the agency?*

III. Resources and staff
12. Is evaluation supported by appropriate financial and staff resources?
13. Does the evaluation unit have a dedicated budget? Is it annual or multiyear? Does the budget cover activities aimed at promoting feedback and use of evaluation and management of evaluation knowledge?
14. Does staff have specific expertise in evaluation, and if not, are training programmes available?
→ *Is there a policy on recruiting consultants, in terms of qualification, impartiality and deontology?*

IV. Evaluation partnerships and capacity building
15. To what extent are beneficiaries involved in the evaluation process?
16. To what extent does the agency rely on local evaluators or, when not possible, on third party evaluators from partner countries?
17. Does the agency engage in partner-led evaluations?
18. Does the unit support training and capacity building programmes in partner countries?

→ *How do partners/beneficiaries/local NGOs perceive the evaluation processes and products promoted by the agency/country examined in terms of quality, independence, objectivity, usefulness and partnership orientation?*

V. Quality
19. How does the evaluation unit ensure the quality of evaluation (including reports and process)?
20. Does the agency have guidelines for the conduct of evaluation, and are these used by relevant stakeholders?
21. Has the agency developed/adopted standards/benchmarks to assess and improve the quality of its evaluation reports?

→ *How is the quality of evaluation products/processes perceived throughout the agency?*

VI. Planning, coordination and harmonisation
22. Does the agency have a multi-year evaluation plan, describing future evaluations according to a defined timetable?
23. How is the evaluation plan developed? Who, within the aid agency, identifies the priorities and how?
24. In DAC members where ODA responsibility is shared among two or more agencies, how is the evaluation function organised?
25. Does the evaluation unit coordinate its evaluation activities with other donors?
26. How are field level evaluation activities coordinated? Is authority for evaluation centralised or decentralised?
27. Does the evaluation unit engage in joint/multi donor evaluations?
28. Does the evaluation unit/aid agency make use of evaluative information coming from other donor organisations?

→ *In what way does the agency assess the effectiveness of its contributions to multilateral organisations? To what extent does it rely on the evaluation systems of multilateral agencies?*

VII. Dissemination, feedback, knowledge management and learning
29. How are evaluation findings disseminated? In addition to reports, are other communication tools used? (Press releases, press conferences, abstracts, annual reports providing a synthesis of findings)?
30. What are the mechanisms in place to ensure feedback of evaluation results to policy makers, operational staff and the general public?
31. What mechanisms are in place to ensure that knowledge from evaluation is accessible to staff and relevant stakeholders?

→ *Is evaluation considered a 'learning tool' by agency staff?*

VIII. Evaluation use
32. Who are the main users of evaluations within and outside the aid agency?
33. Does evaluation respond to the information needs expressed by parliament, audit office, government, and the public?
34. Are there systems in place to ensure the follow up and implementation of evaluation findings and recommendations?
35. How does the aid agency/ministry promote follow up on findings from relevant stakeholders (through e.g. steering groups, advisory panels, and sounding boards)?
36. Are links with decision making processes ensured to promote the use of evaluation in policy formulation?
37. Are there recent examples of major operation and policy changes sparked by evaluation findings and recommendations?

38. Are there examples of how evaluation serves as an accountability mechanism?

→ *What are the perceptions of non-evaluation actors (operation and policy departments, field offices, etc) regarding the usefulness and influence of evaluation?*