

Making Learning Count

Effective Monitoring and
Evaluation of Youth Employment
Programs in the Arab World

TAQEEM INITIATIVE

Making Learning Count:
Effective Monitoring and
Evaluation of Youth Employment
Programs in the Arab World

CONTENTS

Preface	iv
Acknowledgments	v
Acronyms	vi
PART 1: Taqueem and M&E Lessons in Youth Employment	1
1. Introduction	3
2. The Taqueem approach	8
3. Lessons on design, implementation, and use of M&E	12
4. Lessons on M&E capacity development	27
5. Moving forward: Strengthening M&E of youth employment programs	31
References	36
PART 2: Case Studies	39
1. Al-Amal Microfinance Bank, Yemen	41
2. Education for Employment Foundation, Morocco	50
3. Help Leads to Hope, Somalia	57
4. International Youth Foundation, Egypt	65
5. Lothan Youth Achievement Center, Kuwait	72
6. Mennonite Economic Development Associates, Morocco	81
7. SPARK and Birzeit University, Palestine	89

PREFACE

In a region beset by economic and political challenges, creating economic opportunities for Arab youth has long been a leading policy priority. In recent years, international organizations, governments, non-governmental organizations, and social entrepreneurs have responded to this challenge with a wide range of approaches, including interventions focused on skills development, career guidance, job search support, entrepreneurship promotion and various flavors of active labor market programs. In terms of new approaches, it is an exciting time for the region: while more traditional policy approaches are still the mainstay of public responses to the youth employment challenge, there has emerged a range of innovative new approaches, many led by young social and policy entrepreneurs themselves. However, there remains little evidence across the region as to whether such programs have had a positive impact on the economic lives of young people or how to expand successful programs and implement them at the scale needed to resolve the economic challenges facing the region's youth.

The vital need for evidence of what programs work best for youth has led to Silatech's longstanding partnership with the International Labour Organization and the Youth Employment Network in establishing the Taqueem Fund for Evaluation in Youth Employment. Identifying approaches that work requires a significant investment in building evidence about program effectiveness. In working directly with youth-serving organizations through Taqueem, we also have come to realize the challenges faced by smaller organizations in developing and implementing more basic monitoring and program evaluation systems. Towards this end, we are proud of the learning that has come out of the Taqueem Initiative. Moreover, we are excited by the sense of community that has developed among those youth-serving organizations working within the Taqueem Community of Practice, realizing that cooperative approaches to building capacity for monitoring and evaluation has served to foster greater learning and cooperation among such organizations more broadly.

This Compendium serves to document the progress made by the Taqueem Initiative and the organizations that make up its Community of Practice as we move forward in expanding the evidence base for effective youth employment interventions in the Arab region. More importantly, this Compendium is an effort to document our ongoing challenges and the lessons learned in how to better empower local organizations in building their own capacity for monitoring and evaluation. With this in mind, we hope this document serves the broader community working within the youth employment space, helping smaller youth-serving organizations to move forward more effectively in establishing evidence on the effectiveness of their programs, informing donor organizations on how they can more effectively support such organizations, and inspiring governments to better capture learning on the impact of larger scale youth programs. In this way, we can collectively work towards securing a brighter future for the region's young people and build a solid foundation for its economic development.

Nader Kabbani, Ph.D.
Director, Research & Policy
Silatech

ACKNOWLEDGMENTS

This compendium of case studies is the result of a collaborative effort between the ILO, Silatech, and a group of dedicated project staff and consultants associated with the Taqueem project. Under the technical guidance of Drew Gardiner, ILO Evaluation Specialist and Paul Dyer, Silatech Knowledge Program Manager, much of the writing of this report was done by Kevin Hempel, independent consultant for the ILO, World Bank, and other development partners. We would also like to thank the eleven M&E staff and program managers that were part of the Taqueem Community of Practice and helped with the drafting of the case studies: Mohammed Saleh Al-Lai, Al Amal Bank, Hany Amin and Ramy Amin, Better World Foundation, Magda Samy, Ebtessama, Amine Bakkali, Education for Employment Foundation, Abdiaziz Mohamoud, Help Leads for Hope, Rami Shishan, INJAZ, Dan Oliver, International Youth Foundation, Willie Matthew, LOYAC, Khadija Saoudi and Adil Sadoq, MEDA, Nic Van der Jagt, SPARK and Gabool al-Mutawakel, Youth Leadership Development Foundation. We also thank the following members of the Taqueem Technical Team for their dedication to the project and their assistance in drafting the case studies: Chris Wood, Remi Ouellet, and Niklaus Eggenberger. Markus Pilgirm, ILO Manager of the Small Enterprise Unit, provided diligent comments.

Finally, we thank the Jacobs Foundation, Silatech, the World Bank, and the Swedish International Development Agency, for providing resources for Taqueem. This publication was made possible by the NPRP award [NPRP 5-1177-5-174] from the Qatar National Research Fund (a member of The Qatar Foundation).

ACRONYMS

3ie	International Initiative for Impact Evaluation
AMB	Al-Amal Microfinance Bank
CoP	Community of Practice
DIV	Development Innovation Ventures
EFE	Education for Employment Foundation
HLH	Help Leads to Hope
ILO	International Labour Organization
IPA	Innovation for Poverty Action
IYF	International Youth Foundation
J-PAL	Abdul Latif Jameel Poverty Action Lab
LOYAC	Lothan Youth Achievement Center
M&E	Monitoring and Evaluation
MEDA	Mennonite Economic Development Associates
MENA	Middle East and North Africa
NGO	Non-Governmental Organization
RCT	Randomized Controlled Trial
YEF	Youth Entrepreneurship Facility
YEI	Youth Employment Inventory
YEN	Youth Employment Network



PART I

TAQEEM AND M&E LESSONS IN YOUTH EMPLOYMENT



1. INTRODUCTION

Background and motivation

In response to the global youth jobs crisis, government, employers, and labor unions appointed youth employment the central topic of the 101st International Labour Conference in 2012. Entitled “The Youth Employment Crisis: Time for Action,” the conference report reminds the international community that investing in young people is crucial for development. Despite the many national interventions that have focused on providing jobs for youth, the report warns,

[M]ajor gaps in knowledge persist. There have been relatively few rigorous evaluations of all policies and programmes, of their impact in the short and long term, and of their relative cost-benefit, including in developed countries. This needs to be remedied since lessons learned from evaluations can lead to greater programme effectiveness and better targeting of scarce resources. Continuous building of the knowledge base on country policies and programmes and the impact evaluation of the range of measures is a paramount priority (ILO 2012).

The call for more evaluation of youth employment programs began in 2007 when the World Bank published a review of youth employment programs, analyzing data from 300 studies from over 80 countries across the globe (Betcherman et al. 2007). This review concluded that the overall evaluation evidence of youth employment programs was weak, especially in some developing regions. A case in point: the authors found only two robust evaluations of youth employment programs from sub-Saharan Africa and none from the Middle East and North Africa (MENA) region. The review’s major finding was that the evidence base on which governments and development partners can make informed programming decisions to help young people transition from school to work was extremely limited and that major improvements have to be made.

Since 2007, the policy imperative for large-scale solutions that help youth enter the world of work has only grown. Many observers talk of a “global youth employment crisis,” and the International Labour Organization (ILO) has warned about the “scarring” effects of lasting youth unemployment and inactivity (ILO 2012). Outcomes for youth in the Arab world are of particular concern. With the region’s youth population peaking as a share of total population, youth across the region have experienced persistent, high rates of unemployment and underemployment. Overall, unemployment rates for youth in this region are nearly twice

those of the international average (ILO 2013). Moreover, the need to put young people to work has become all the more important since the tumultuous events of the Arab Spring, when millions of Arab youth took to the streets to voice their frustrations with the economic opportunities available to them.

Our response

In response to the lack of evidence behind youth employment interventions in the Arab world, the ILO joined with partners, including Silatech, the World Bank, the Jacobs Foundation and the Swedish International Development Foundation in 2009 to establish the Taqueem Initiative. The objective of Taqueem is to engage with youth employment policy-makers and practitioners and support them in enhancing the monitoring and evaluation (M&E) of their programs (see box 1).¹ Through Taqueem, the partnership has responded to a niche demand among youth-serving organizations (YSOs) to provide technical assistance on M&E which is appropriate for their size and resources while, at the same time, creating learning opportunities that provide the broader youth employment community with better knowledge on what works in generating youth employment opportunities.

BOX 1. Defining terms

Monitoring: A continuous process of collecting and analyzing information to see how well a project, program, or policy is being executed and performing against expected results.

Evaluation: A systematic, objective assessment of an ongoing or completed project design, implementation, and result to determine its relevance and the fulfillment of objectives, efficiency, effectiveness, impact, and sustainability.

Impact evaluation: A special type of evaluation that assesses the changes in the wellbeing of individuals, households, or communities that can be attributed to a particular intervention.

Source: Hempel and Fiala (2011)

Since 2009, the Taqueem Initiative has engaged with over 100 YSOs working across the Arab region, engaging representatives of these organizations in initial capacity-building clinics focused on providing basic training in monitoring and evaluation methodologies. The Initiative has also provided more targeted and in-depth support to twelve organizations focused on the development of individualized M&E strategies. These include programs running in Egypt, Jordan, Kuwait, Morocco, Palestine, Somalia and Yemen. This has included two ongoing impact evaluations (Bamyan Media's El-Mashrou3 program and MEDA Maroc's 100 Hours to Success). Building on its work with Taqueem, Al-Amal Microfinance Bank has also been working with Silatech and scholars from Georgetown University on developing an impact evaluation. Details on each of these organizations and evaluation support provided to them is provided in table 1 below.

¹ The Taqueem Initiative was originally established by the Youth Employment Network (YEN). YEN was created in 2001 as a partnership between the United Nations, the ILO, and the World Bank to mobilize action on the commitment of the Millennium Summit for decent and productive work for young people. While the YEN partnership was concluded in 2014, the activities initiated under YEN in support of Taqueem and M&E more broadly continue through the ILO's Youth Employment Programme.

Purpose of this volume

Effective monitoring and evaluation has both external and internal uses. Internally, it can help an organization improve its results by tracking successes and failures and making adjustments in programming based on lessons learned. Externally, public evidence on what works and what does not in designing and implementing interventions can improve outcomes on a global level. In this regard, the two objectives of this compendium are to (1) contribute to the knowledge and learning around “what works” in youth employment and how organizations can contribute to this knowledge base and (2) to provide important insight on how the knowledge base can be expanded by strengthening learning and accountability across organizations and programs, large and small.

The organizations supported through Taqueem represent a wide range of youth employment interventions. These include programs aimed at vocational and life-skills training; on-the-job training and internships; employment services and job-placement support; entrepreneurship education; youth microfinance; and, in many cases, combined approaches. On the one hand, this diversity provides a wealth of potential learning from a range of different types of interventions. On the other, this diversity makes it more challenging to draw generalizable lessons from across the supported programs. Moreover, of the twelve programs with which Taqueem worked most closely, only two are implementing rigorous impact evaluation approaches that provide robust evidence of program effectiveness, and none of these were completed by the time of writing.

As such, it remains too early to focus on program outcomes and impact. Rather, this volume focuses on lessons learned from efforts to provide YSOs the capacity-building and technical assistance needed for building the foundations of a long-term effort to increase our collective understanding of the impact of such programs. To this end, we provide examples of how the organizations we have supported have improved their monitoring and evaluation systems, as well as identifying challenges and lessons learned from attempts to implement rigorous impact evaluations. The volume also provides recommendations for development partners on how they can further efforts to support results-based learning.

Our overarching goal is to encourage future learning in the area of youth employment and to increase the recognition that this collective effort is only possible with targeted M&E support to implementing organizations aimed at helping them overcome technical and financial barriers. We believe that strengthening capacity for monitoring, evaluation, and impact evaluation across youth employment practitioners is a key priority in order to foster knowledge generation and learning that can inform future policy and practice. This is especially the case in the Arab world, where the challenge of providing young people with meaningful employment opportunities is so crucial.

The compendium is divided into two parts. Part I has five sections. Following from section 1, Introduction, Section 2 provides an overview of the Taqueem approach to supporting organizations with training, capacity-building, technical assistance and funding to promote learning through the application of monitoring and evaluation. Section 3 draws on the experience of supported organizations in applying M&E knowledge, identifying common challenges to implementation of M&E strategies and providing recommendations for implementing organizations on putting such strategies into practice. Section 4 focuses on lessons for development partners on how M&E training and capacity-building interventions can be best designed to empower local organizations. Lastly, section 5 details priorities for the development community, particularly those addressing youth employment issues forward and how Taqueem is positioning itself to aid in those efforts.

Part 2 is a collection of seven case studies which serve as a foundation for the compendium providing practical illustrations of how organizations are meeting the challenges of M&E. Table 1 outlines the case study topics and the organizations supported under the Taqueem Community of Practice (CoP).

TABLE 1. Overview of Taqueem-supported organizations

Case study #	Organization	Country	Program name	Program description	Type of M&E conducted
Case studies					
1	Al-Amal Microfinance Bank (AMB)	Yemen	Youth Loan Fund	<ul style="list-style-type: none"> Start-up and business expansion loans Entrepreneurship support (training, mentoring, incubation) 	Data collection instruments for outcome monitoring Impact evaluation design
2	Education for Employment Foundation (EFE)	Morocco	Finding a Job is a Job (FIJ)	<ul style="list-style-type: none"> Job-search skills training 	Pre-/Post-evaluation (non-experimental)
3	Help Leads to Hope (HLH)	Somalia	Fishing Sector Livelihood Program for Youth	<ul style="list-style-type: none"> Technical training Financial education Bank account In-kind grant (equipment) 	Data collection instruments for outcome monitoring
4	International Youth Foundation (IYF)	Egypt	Egypt@Work	<ul style="list-style-type: none"> Employment services Business support Internship/apprenticeship 	Cloud-based data management system Impact evaluation design
5	Lothan Youth Achievement Center (LOYAC)	Kuwait	Summer Program	<ul style="list-style-type: none"> Internship placement Group projects Volunteering 	Cloud-based data management system
6	MEDA Maroc	Morocco	100 Hours to Success	<ul style="list-style-type: none"> Training in life skills, entrepreneurship, financial education Internships Savings account 	Impact Evaluation
7	SPARK and Birzeit University	Palestine	Strengthening technology entrepreneurship for a sustainable Palestinian knowledge-based economy	<ul style="list-style-type: none"> University-based entrepreneurship education Coaching Matching with financial institutions 	Cloud-based data management system Impact evaluation design

(continued)

TABLE 1. (continued)

Case study #	Organization	Country	Program name	Program description	Type of M&E conducted
Supported organizations, no case study					
1	Bamyan Media	Egypt	El-Mashrou3	<ul style="list-style-type: none"> ■ Televised entrepreneurship competition ■ Information on support services ■ Online training and other resources 	Impact Evaluation
2	Better World Foundation	Egypt	Path to Your Career	<ul style="list-style-type: none"> ■ English language classes ■ ITC classes ■ Life-skills training 	Monitoring system
3	Ebtessema Foundation	Egypt	A Right for Equal Life: Young Adults with Disabilities	<ul style="list-style-type: none"> ■ Social & life-skills training ■ On-the-job training and coaching ■ Job placement 	Monitoring system
4	INJAZ	Jordan	Company Program	<ul style="list-style-type: none"> ■ Entrepreneurship education 	Monitoring system
5	Youth Leadership Development Foundation	Yemen	KHADIJA	<ul style="list-style-type: none"> ■ Women's entrepreneurship training ■ Business grants ■ Mentoring ■ Business development services 	Monitoring system



2. THE TAQEEM APPROACH

With Taqueem, we have developed an approach to M&E that works closely with YSOs to provide initial training on M&E, followed by close cooperation with supported organizations as they develop and implement program-relevant M&E systems. The Taqueem support model is built on the firm belief that generating a critical mass of knowledge on youth employment requires a multisectorial effort. It requires tapping into the wealth of programming carried out by non-governmental organizations (NGOs) and other implementing organizations. In that, the challenge consists of building the right learning systems and tools so that practitioners can rigorously test the effectiveness of their interventions and programming innovations, systematize their tacit knowledge, and contribute to a larger body of evidence that is widely available and that can be built on by other practitioners in the future.

There is a tremendous demand among youth-serving organizations and NGOs more broadly for technical advice on evaluation, driven largely by growing demands by donors and other development partners for evaluation results. At the same time, the provision of assistance in improving results measurement are in short supply. Moreover, generating more and better evidence on youth employment is intrinsically linked to building the capacities and skills of individuals and organizations to set up strong M&E systems and to carry out robust impact evaluations.

The general approach to capacity-building on M&E used within Taqueem has evolved over time. The Taqueem Initiative is itself rooted in the Youth Employment Network's support for M&E in sub-Saharan Africa which began in 2007. Initially, efforts focused on identifying opportunities for conducting impact evaluations. YEN identified organizations in need of support and held trainings for them called impact evaluation clinics. The outcomes of these trainings led to several impact evaluations within the sub-Saharan Africa region, supported by YEN's Youth Entrepreneurship Facility. However, resource investment into this approach was high and, more importantly, only a few organizations were able to benefit from the support. In developing the approach for Taqueem, we developed a suite of capacity-building services complemented by targeted financial resources. In this way, we were better able to tailor services to the needs and capacities of participating organizations. For some, this meant focusing more on basic monitoring rather than impact evaluation, building the foundations for improved outcome measurement over time.

Taqeem's support to YSOs is delivered through three iterative steps:

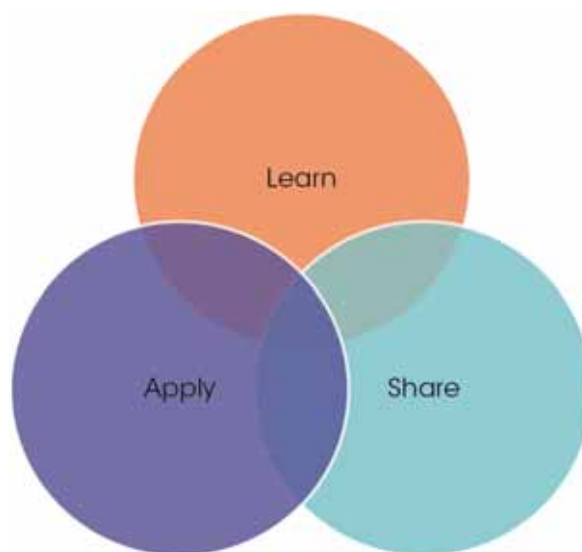


TABLE 2. Sample agenda of a Taqeeem evaluation clinic

Day 1	Day 2	Day 3	Day 4
Welcome and introductions	Data collection tools	Impact evaluation methods	Data analysis
Why M&E and impact evaluation	Group work	Fictional case study: Identification of a counterfactual	Group work: finalizing evaluation plans
Project marketplace: case studies introduce projects	IT-based data collection (web, SMS, etc.)	Group work	Group presentations of evaluation plans
Refresher on theory of change, indicators, and evaluation questions	Survey design	Data for impact evaluation: sampling, collection, management	
Group work	Group work	Group work: finalizing the concept notes	

- Learn:** Learning is provided through M&E training for practitioners, primarily through evaluation clinics, and complemented with “webinars” on specific topics of interest. Typical training clinics last 3–4 days and cover the basics of M&E, impact evaluation methods, and data collection tools (see table 2). The curriculum of Taqeeem’s capacity-building builds greatly on the handbook *Measuring Success of Youth Livelihood Interventions: A Practical Guide to Monitoring and Evaluation*, available in English, French, and Arabic (Hempel and Fiala 2011). During the clinics, a selected number of participants serve as live case studies whose options for stronger M&E or impact evaluation are discussed and outlined during group work sessions. Training programs also serve as an initial opportunity to match the needs of supported organizations with outside experts.

TABLE 3. Overview of M&E support provided by Taqueem

Strengthening M&E systems	Preparing for impact evaluation	Conducting impact evaluation
<ul style="list-style-type: none"> ■ Refinement of indicators at the output and outcome levels ■ Development and/or refinement of data collection strategy comprising easily applicable instruments ■ Development of scorecards and reporting templates ■ Testing and finalizing of instruments ■ Determining sample populations ■ Training staff on data collection ■ Support development of M&E data management system and/or database ■ Data analysis and report writing 	<ul style="list-style-type: none"> ■ Refinement of indicators at the outcome and impact level ■ Preparation of a time line and budget of the evaluation ■ Defining the sample size, the selection method and carrying out power calculations ■ Setting up the evaluation team: recruitment of researchers and investigators ■ Developing concept notes for evaluation ■ Helping to secure funding for the evaluation ■ Developing and piloting the survey instrument 	<ul style="list-style-type: none"> ■ Finalizing concept notes ■ Finalizing sample and selection method ■ Support for training field work team in data collection ■ Serving on research teams as research assistant, principle investigator or research coordinator ■ Assistance in conducting baseline survey ■ Assistance in tracking survey participants ■ Assistance in conducting follow-up survey ■ Data analysis and report writing

- **Apply:** Classroom-based training is complemented with “learning by doing” acquired during the implementation stage. Following initial training, organizations devise new and improved monitoring, evaluation and impact evaluation plans for their organizations. During this stage, participants receive coaching by technical experts. Initially, assistance focuses on developing an appropriate M&E or impact evaluation plan based on the needs of the organization. Subsequently, coaches help with the implementation of tools and approaches, including through field visits. Small grants are provided to cover costs for critical expenditures to help the organizations implement their plans. In the specific context of impact evaluations, Taqueem experts have served on the study teams as either principal investigator, research associate or study manager, and helped organizations secure outside funding. An overview of the types of follow-up support services provided are listed in table 3.
- **Share:** Knowledge-sharing is achieved through the dissemination of results and lessons learned from the M&E systems and impact evaluations. Towards this end, supported organizations join a community of practice (CoP) to facilitate knowledge exchange and ensure continued engagement in building the capacity of member practitioners to conduct evaluations of their program outcomes, including impact evaluations. Members of the CoP benefit from regular interactions with researchers and evaluation experts, and meet regularly at workshops and youth employment-related events to share experiences. Beyond the CoP activities, the ILO administers an online networking site, which includes members of the CoP within a larger network of youth employment practitioners. The networking site is used to facilitate interaction and to share relevant M&E and impact evaluation documents (see box 2).

BOX 2. ILO's Fund for Evaluation networking site

Created in 2011, ILO's online networking site grew to a membership of approximately 900 youth employment practitioners by 2013. This member-driven website:

1. Provides information related to evaluation activities offered by ILO and its partners (call for proposals, application forms, learning events)
2. Connects researchers, the policy-makers and practitioners around the topic of youth employment and evaluation.
3. Provides an outlet to share training materials, results, conclusions and lessons learned from impact evaluations and M&E supported by ILO.

To get more information and become a member, visit <http://yenclinic.groupsites.com>

The shared goal of these efforts has been to generate new knowledge on what works in youth employment, to strengthen the learning systems of implementing organizations, and to develop local M&E and impact evaluation capacity – that is, building a critical mass of professional evaluation staff in these regions. Beyond this, however, the Taqueem Initiative is determined to become a source of learning for the broader youth-serving community. Dissemination of evaluation evidence has been transmitted regularly to policy-makers through ILO member states. Findings and recommendations have also been extracted and synthesized through web-based forums, including the ILO's Youth Employment Inventory (<http://www.youth-employment-inventory.org/>). This document and the accompanying case studies represent an effort to expand our dissemination to a wider audience of stakeholders in the youth employment space and the development community more broadly.



3. LESSONS ON DESIGN, IMPLEMENTATION, AND USE OF M&E

Our experience in working with YSOs to develop and implement more effective M&E systems has highlighted a number of core challenges faced by M&E practitioners. Building on this experience, we offer advice and lessons learned from efforts to meet those challenges, supported by examples from the experience of Taqueem-supported organizations. These lessons have been collected through consultations with M&E officers, program managers, researchers and other technical staff who have received technical advice, training and financial assistance in the design, implementation and usage of monitoring and evaluation systems.

Lessons are presented according to the four steps of the basic M&E cycle (figure 1). This basic model is intended to be generic and could apply to any project or a program, with the goal being to improve program results through a systematic and iterative feedback loop between planning and practice.

Planning and design of M&E

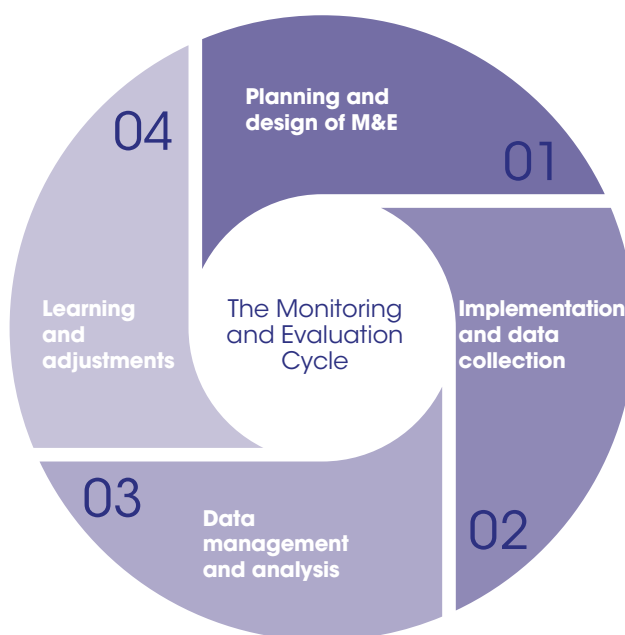
The M&E cycle can be thought of as a parallel process to the project cycle. In the same way that good planning and project design is crucial for an intervention to be successful, the planning and design of an M&E system is important to ensure that learning can take place. Rather than making M&E an afterthought, ignored until reporting requirements arise, M&E should be used strategically and designed in parallel to the intervention in a way that meets the needs of the implementing organizations.

When setting up an M&E system, one needs to be conscious of the dual role that M&E can play. In the short term, often referred to as the monitoring function, the M&E system is critically linked to project operations, providing real-time learning support to those involved in implementation. In the longer term, more often associated with the evaluation function, the M&E system focuses on higher-level, impact-related learning where findings and recommendations only emerge after longer periods and more in-depth assessments of whether a project is working or not.

Securing buy-in from senior management and other stakeholders

Strengthening M&E and carrying out quality evaluations is difficult without an enabling environment, and securing buy-in from the organization's senior management is essential to ensuring that this environment is put in place. Senior management is responsible for an organization's overall policy and direction, which includes performance measurement and accountability to donors. They are also responsible for allocating the required resources in terms of staff time and funding. While building internal ownership is crucial, so is getting buy-in from external stakeholders, including

FIGURE 1. The monitoring and evaluation cycle



board members, key partner organizations, and donors, since their influence on the organization may directly affect the priority given to monitoring, evaluation, and learning.

The importance of management ownership for M&E has been apparent in the success or failure of Taqueem grantees when constructing their results measurement systems. At Al-Amal Microfinance Bank, for example, there was significant support from the Bank's CEO, who had participated in the evaluation clinics (see Case Study 1). In this case, part of the buy-in by the Bank's board and management was facilitated by external partnerships which provided funding in support of the bank's M&E plans. In contrast, at LOYAC, the Board was not fully aware of the potential benefits of strong M&E to the organization and restricted expenditures in this area.

It is recommended to identify a highly placed "champion" within the organization who understands the importance of M&E, will provide support and shield attacks from any counter reformers or those contesting the costs and time spent on implementing an M&E system. Organizations should also understand the priorities of key stakeholders (e.g. related to accountability, knowledge generation, fundraising, etc.) and present the benefits of a strong M&E system in the light of meeting their respective needs. Seek out their evaluation questions so your organization's M&E system can be designed to meet the different interests and demands.

M&E organizational structure, staffing, and responsibilities

Quality M&E demands considerable time and resources in order to plan, collect data and report. How should the M&E function be set up within the organization and what should be the role of M&E? In the experience of organizations working within Taqueem, some faced high turnover in the M&E role, in part because qualified M&E officers are in short supply and thus often highly sought after by other organizations. Many organizations combined the M&E function with programming or communications roles, and in these cases, the M&E function lost out to other responsibilities. Ideally, a monitoring, evaluation and learning unit is set up as a crosscutting unit within an organization, placed between the technical/programming units and the senior management in the organization's hierarchy.

Experiences from Taqueem organizations has proven that while one staff member within an organization should be fully dedicated to M&E, this dedicated M&E officer cannot operate in isolation and needs to be fully integrated into planning and other processes. Without this integration, the important feedback loop between programming and M&E cannot take place. This can include providing job descriptions for all employees and volunteers that include responsibilities for M&E that are relevant to their roles. Important examples can be found in the cases of Education for Employment in Morocco (Case Study 2) and LOYAC in Kuwait (Case Study 5).

Enhancing staff capacity

When our partners have struggled with implementation of their M&E systems, the most common reason was a lack of M&E knowledge and experience amongst the wider staff. The M&E function requires a high level of technical knowledge with some background in statistics and econometrics. However, M&E responsibilities do not fall only to the M&E team, but extend to all management and staff. As such, all relevant staff require some basic technical understanding of M&E theory and practice. In some cases, M&E teams might include external specialists, who can provide needed technical knowledge at strategic times. For instance, MEDA Maroc chose to work with external researchers because their evaluations required high levels of technical capacity for data analysis and reporting (see Case Study 6). The International Youth Foundation worked with consultants specializing in cloud-based M&E systems (see Case Study 4).

In this regard, it is recommended that to have an M&E “guru” on staff, a senior specialist to whom staff can go with questions and who can mentor the rest of the team (see box 3). It is also advised to have staff participate in M&E training while recognizing that training is not a “silver bullet” and should be complemented by on-the-job learning and coaching. Finally, each staff member must understand their role in the M&E system and why it is important to the organization. It is necessary to map out the “data flows” within an organization and to determine who is responsible for each flow.

Establishing effective partnerships

A good evaluation always involves the establishment of effective partnerships between implementing organizations and external collaborators. Experienced evaluation researchers bring critical skills in evaluation design, research methodologies, data

BOX 3. Skills to look for in an M&E manager

- Good understanding of youth, decent work and poverty issues
- Understanding of quantitative and qualitative tools for data collection (surveys, focus groups, etc.)
- Ability to check and clean data
- Ability to analyze data
- Report-writing and analytical skills
- Facilitation, teaching, and reflective learning skills
- Excellent communication and presentation skills
- Proficiency in Microsoft Office and other statistical software as needed

BOX 4. Capacity development at INJAZ Jordan

INJAZ Jordan is a good example of an organization investing heavily in M&E capacity development. Over the period 2011–13, individuals within the organization completed the following learning opportunities:

- Participated in two ILO evaluation clinics in Amman and Doha, which provided participation with basic understanding of M&E design, research methodologies, and data collection
- Organized an internal M&E training with their expert consultant, providing staff with a background in M&E work plan development, indicators, and participatory M&E
- Sent their M&E manager to attend JPAL's executive education course in impact evaluation, which provided a deeper understanding of randomized controlled trials, the use of statistics software like STATA/SPSS, the application of power calculations and regressions, and the ability to formulate research questions correctly
- Shared knowledge with partner institutions, providing advice on impact evaluation strategies
- Redrafted terms of reference for all M&E staff to ensure that M&E responsibilities remained priorities for all

collection and analysis, and often have good links to public authorities for policy outreach. Working with external collaborators also ensures impartiality of results and brings new approaches to an evaluation that may not be immediately obvious to those operating from within a program.

The potential synergies go well beyond the implementing agency's input into the evaluation questions. MEDA's "100 Hours to Success" study is a rare case of an implementing organization being actively involved in the impact evaluation process (see Case Study 6). In this particular case, the comparative advantage of having the implementing organization actively participate in the study included MEDA's support in aligning the data collection cycle to the programming cycle, fostering take-up, and leveraging beneficiaries' trust in the organization, which was especially helpful when collecting sensitive information.

Despite the benefits of collaboration, partnerships between implementing organizations and researchers often fail because research priorities and methodologies do not match with realities of program operations and because organizations are not clear on specific deliverables before hiring the researchers. With these concerns in mind, the following recommendations apply:

- Implementing organizations should explore links with local or international research institutions and universities by looking at their stated research priorities and the research interests of faculty and staff to ensure alignment with the interests of the organization. Research symposiums and conferences also provide opportunities to identify potential collaborators.
- Organizations should work closely with external collaborators to ensure that the latter understand the details of program design and implementation, assuring that proposed methodologies for evaluation are in alignment with realities on the ground.
- Once a team of collaborators is identified, organizations should meet with researchers regularly, keeping them up to date on any changes in program design. This allows the evaluation team to adjust research design accordingly.

Developing a clear results chain

A solid results chain is the basis for effective M&E and impact evaluation. Whether one is designing a basic M&E system or a sophisticated randomized controlled trial (RCT) evaluation, one should start with determining a program's results chain (also sometimes known as theory of change, results framework, or logical framework). The results chain is the logical outline of the how the intervention is expected to lead to the desired results, explaining the process of change by outlining short-, medium- and long-term outcomes of an intervention. For impact evaluations, it helps a research team define learning or research objectives and corresponding outcome indicators.

This is why developing and agreeing on a fleshed-out results chain was the first step in the support to youth-serving organizations under the Tajeem Initiative. For each of the supported organizations, results chains provided the basis for articulating learning objectives and identifying relevant indicators, while helping the teams to identify underlying assumptions. Staff and other key stakeholders should be involved in the process, in order to build a mutual understanding of the intervention.

Learning objectives as a starting point

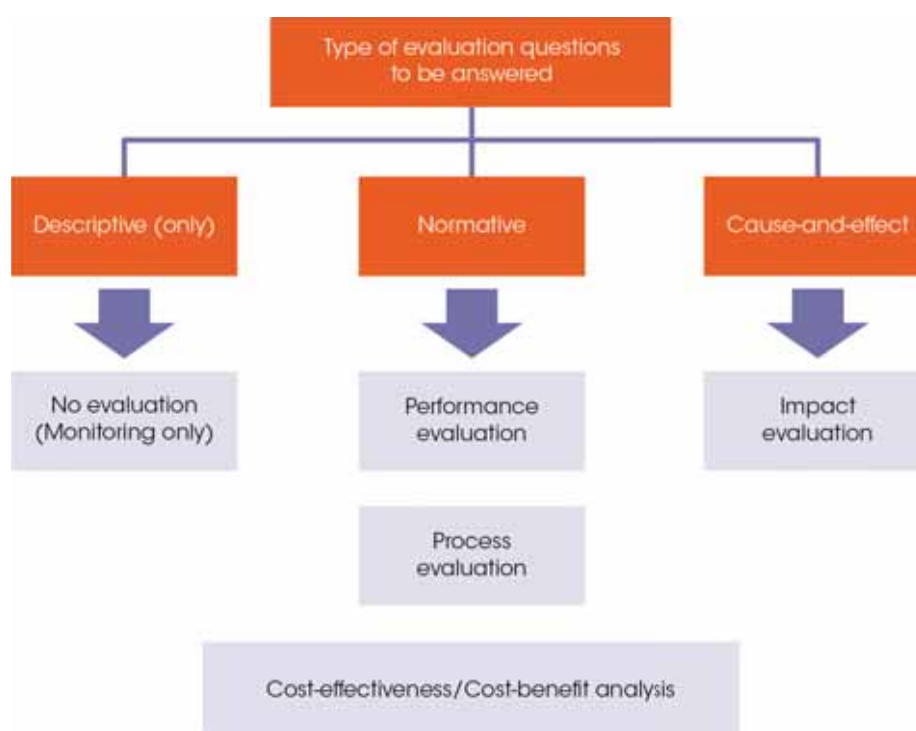
As an organization begins choosing between different monitoring, evaluation and impact evaluation approaches, it should begin with its objectives in regard to what it wants to learn about its program. There are usually two main objectives of an M&E approach: learning (for program management and broader knowledge generation) and demonstrating legitimacy (for accountability to funders and sustainability of the organization). Even though donor demands often trigger an organization's "desire" to strengthen their M&E capacity, learning is ultimately the more important objective, since it goes beyond merely reporting on activities and targets and aims at understanding why the program achieved what it did (or did not), and how it can do better. Therefore, M&E needs to support programming, not vice versa.

In this regard, it is important for organizations to realize that monitoring, evaluation, and impact evaluation serve different needs. Monitoring answers descriptive questions (how the project is being implemented, how many beneficiaries are being reached, whether participants are satisfied, etc.). Traditional evaluations such as performance evaluations, process evaluations or cost-benefit analyses are intended to answer normative questions; that is, comparing what is taking place to what should be taking place (e.g., has the project performed as expected?). Impact evaluations, in turn, answer cause-and-effect questions; that is, whether the effects observed have been achieved as a result of our intervention. None of them is better or worse than the others – they just serve different objectives (see figure 2).

Any M&E system should be grounded in the learning objectives and evaluation questions of the organization. If external researchers are involved, they must recognize that each evaluation should make the organization's learning objectives a first priority. Sometimes this may require the researcher to also be a facilitator, helping the organization identify and prioritize its learning needs.

Matching evaluation methods to program maturity

The appropriate M&E strategy has to be based on the maturity and capacity of the organization and program. In the case of our support to youth employment practitioners in the Middle East and North Africa, this translated into a differentiated approach. Larger and more established

FIGURE 2. From evaluation questions to evaluation design

Source: Hempel and Fiala (2011)

programs pursued impact evaluations (e.g., MEDA Maroc) or focused on a better management and analysis of incoming outcome data (International Youth Foundation), while many others emphasized establishing or improving their basic M&E systems and establishing data collection procedures (e.g., SPARK, LOYAC, Help Leads to Hope).

Many organizations need to start small and keep their M&E efforts simple at first. Even when organizations did not pursue an impact evaluation, there was a tendency among organizations supported by Taqueem to want to launch many new tools (surveys, focus groups, ICT-based data collection, participatory tools, etc.) at the same time. This strategy proved too ambitious and resource-intensive for many. As a result, several organizations (e.g., Al-Amal Microfinance Bank, LOYAC, Help Leads to Hope) ended up dropping planned data collection efforts over time, instead focusing only on the most important ones.

A “light” M&E system is often more appropriate for smaller operations, especially when staffing for M&E is limited. This will facilitate data collection, make analysis easier and produce simple reports with little strain on delivery processes. Over time, and as data collection capacity increases, these tools can be expanded and improved to yield more detailed information.

Planning for research-oriented impact evaluations

Monitoring and evaluation strategies in youth employment are increasingly being initiated and led by external consultants, researchers and academics, in collaboration with implementing organizations. Researchers often advocate for impact evaluation designs using randomized

approaches. As we will see from several Taqueem-supported examples, appropriate planning and design is required in order for these approaches to be successful.

First, prospective evaluations, those that are planned at the same time as the program itself, provide much more leeway in terms of methodologies, and make for a stronger evaluation design. In most cases, a baseline survey of potential participants is required, which needs to be implemented before the intervention to be evaluated starts. As such, planning for an evaluation must begin at a very early stage, in line with plans for rolling out the program itself.

At the same time, it is usually advisable to start with basic M&E, rather than investing in a full impact evaluation study, during the program set-up stage, until implementation is smooth, basic beneficiary satisfaction is established, and the program meets its participation targets. Once introductory challenges related to program rollout are ironed out, impact evaluation can be carried out during a subsequent pilot phase. The risks of not doing this are highlighted in MEDA Maroc's experience with evaluating the impact of its 100 Hours to Success program in Morocco (see Case Study 6). In this evaluation, the organization faced a risk of bias in their evaluation results because available monitoring data failed to provide an accurate estimate of the take-up rate of training participation. After the study started, poor program take-up required revisions to the study that introduced the risk of selection bias, frustrating the ability of the evaluators to accurately assess the impact of the program on beneficiaries.



“The methodological gold standard here is appropriateness, not any one particular method.”

(Patton 2008a)

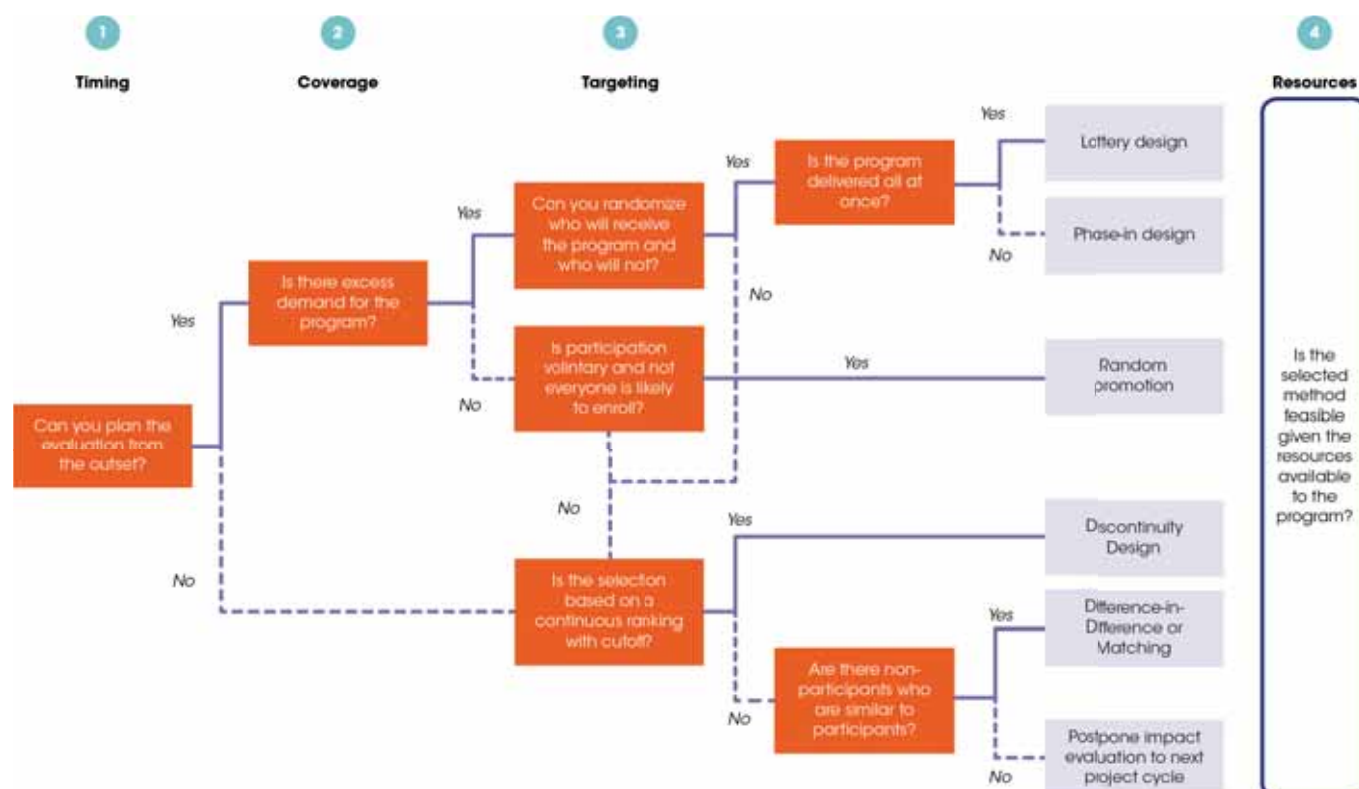
There is a long menu of research methodologies for evaluation teams to choose from, both experimental and non-experimental. While experimental methods are considered by some to be the most “rigorous” method, the best evaluation method is the one that best fits the design features program. Factors such as the stage

of program implementation, program coverage, targeting procedures, and available resources determine which design is the most appropriate (see figure 3). For more information on particular impact evaluation methodologies, readers are referred to Hempel and Fiala (2011).

A case in point is Al-Amal Microfinance Bank in Yemen. Originally, evaluators had planned an RCT to assess the impact of lending and other financial services for youth on employment creation, business expansion, and other outcomes. Yet, senior management of the bank made it clear that such a design would be unacceptable in the context of a recently founded financial institution that cannot afford to exclude potential clients for the purpose of an evaluation. While Taqueem's technical assistance focused on strengthening the Bank's basic M&E system, the evaluation team designed a randomized promotion evaluation that was more suitable for an intervention with universal coverage. The baseline for the evaluation was initiated in 2015. Similarly, IYF in Jordan explored several options for an appropriate impact evaluation design, though none was deemed feasible in the end (see box 5).

Implementation and data collection

While good planning and design of M&E is fundamental to being able to generate relevant knowledge and learning, the bulk of the M&E work and resources is required during the data

FIGURE 3. Decision tree for choosing impact evaluation techniques

Source: Hempel and Fiala (2011)

collection stage. Good data collection and the ability to transform data into information and knowledge is the key to successful project implementation.

Deciding which data collection methods to use

Taqeem-supported organizations have used a wide variety of data collection instruments, ranging from individual surveys to focus groups and key informant interviews to more innovative techniques like SMS surveys and GPS tracking (table 4). Help Leads to Hope, LOYAC, and Al-Amal Microfinance Bank rolled out new quantitative tools that included baseline and follow-up surveys, while EFE Maroc, IYF, and SPARK all incorporated focus groups or other qualitative approaches (e.g., significant change stories) into their M&E systems.

The choice of data collection methods should be determined by the type of information needed, which is related to the success indicators that have been established in the project strategy. Data collection methods are either quantitative or qualitative, and often involve a combination of both. The population coverage of an organization's data collection is key, and, in most cases, a sample of the targeted population will need to be determined based on a sampling size and survey frame. Importantly, no matter the technique chosen, data collection is resource intensive. Are the approaches to be selected feasible, appropriate, relevant, cost effective and timely?

BOX 5. Matching project design to evaluation design for IYF's Youth:Work Jordan Project

IYF Jordan worked with Taqueem to design an impact evaluation for their Youth:Work project, an intervention intended to provide new opportunities for youth in marginalized, poor communities, enabling them to be self-employed in microenterprises. Activities under the intervention include business training, mentoring and access to microcredit.

In considering an approach to evaluating the impact of the program on beneficiaries, IYF's first choice was to conduct an RCT. After considering the approach, this was not deemed feasible as IYF did not expect that there would be over-subscription of interested participants, normally a precondition of an experimental evaluation. Moreover, IYF has strict criteria and eligibility requirements for its trainees, and did not want to randomize selection of participants.

Instead, a quasi-experimental design called a regression discontinuity design (RDD) was proposed to evaluate the project. Working from IYF Jordan's existing screening process, a scaled scoring instrument would be applied that would allow IYF to quantify its applicant review process. This scoring instrument would be used to set the discontinuity for the RDD. A specific score on the scale would be set that would determine which individuals got sorted into the treatment (participants) group and which got sorted into the control group.

When experimental designs are not possible in an evaluation, the RDD allows evaluators to compare results between participants and non-participants on either side of the discontinuity in order to estimate the local treatment effect. This estimation is possible because these marginal participants and marginal non-participants will differ only slightly in underlying characteristics, thus reducing the bias that would otherwise be included in comparing the whole body of participants and non-participants to each other.

In the end, IYF decided to opt out of an impact evaluation, as it became clear that a number of factors would prevent accuracy and reliability of evaluation results:

- Only 350 students were expected to participate in the training over the two years of the project. RDDs by nature require much higher relative sample sizes to secure statistical significance of results than RCTs. With the small numbers of participants, any required sample is already a significant share of participants.
- With all RDDs, there is a risk that the discontinuity will not hold firm. That is, as the program is implemented, program managers may feel increasing pressure to admit more participants (or may begin being more restrictive). In this case, the program would provide services to a number of those deemed ineligible by the score, thus compromising the study.
- An added complication was that training activities were outsourced to contracted service providers that faced their own time and budget constraints. The added M&E tasks to complete registration and exit forms related to the RDD approach would have placed too much of a burden on these teams, and detracted from their core work on training promised to the donor.

In this regard, it is recommended to leverage and improve existing data and tools where possible, instead of creating new ones. Many youth employment organizations have a limited set of resources. In an effort to improve a monitoring system, there can be a tendency to introduce new tools to collect data. Instead, internal M&E staff should try to build on and improve existing data collection tools. Herein, Al-Amal Microfinance Bank provides a good example (Case Study 1). The bank already had in place many data collection and client follow-up procedures through its standard loan management processes. Instead of creating new data collection tools, the bank adjusted its loan application form to capture higher-quality data. Similarly, tablets were introduced to loan officers, allowing them to improve upon paper-based data collection and reduce time and error.

For organizations seeking to identify appropriate data collection instruments, it is important to understand how accurate measures need to be. Overall, in quantitative approaches, closed

TABLE 4. Data collection tools used by supported organizations

Organization	Baseline survey	Follow-up survey	Tracking survey	Special interest survey	Skills test	Focus groups	Employer survey	Most significant change
MEDA Maroc	✓	✓	✓			✓		
Education for Employment Foundation	✓	✓				✓		
International Youth Foundation	✓	✓			✓	✓	✓	
SPARK and Birzeit University	✓	✓						✓
LOYAC	✓	✓			✓		✓	
Al-Amal Microfinance Bank	✓	✓		✓		✓		
Help Leads to Hope	✓	✓			✓	✓		

questions and face-to-face surveys are the most accurate, but have contextual limitations. Mixed-method designs – that is, those that explicitly adopt both quantitative and qualitative methods in data collection – provide the strongest approach. The advantages of a mixed-method evaluation approach are by now well established (Adato 2011; Bamberger, Rao, and Woolcock 2010; Leeuw and Vaessen 2009).

Gathering data from the field

When preparing for data collection, it is important to carefully choose fieldworkers and train them in the selected methods and questions. With this in mind, there are several important recommendations that can be made based on the experience of Taqueem-supported organizations.

First, in preparing for field work, it is important to pilot data collection instruments before they are deployed. Testing survey questionnaires and other data collection tools is crucial to ensuring that the formatting and phrasing of questions are understandable and to confirm that the surveys do not take too much time to complete. Questionnaires should be tested on a selection of individuals similar to those in the program. Ideally, these will include program beneficiaries; however, it is important that results from these pilot surveys are not included among the final respondents so that respondents are not influenced by previous exposure. Difficulties that arise in neglecting testing are seen in the example of Help Leads to Hope in Somalia (Case Study 3), where most of the beneficiaries refused to answer questions during the follow-up survey because they did not feel comfortable providing information directly related to their personal income, mental health, or family affairs.

Second, good preparation for data collection requires training and preparation for the enumerators. For each tool used, each question should be explained with the field team, ensuring

that enumerators understand the questions and the intent behind them. Careful instructions should be given to enumerators on how to ask each question in order to obtain useful information.

Attention to the proper timing of follow-up surveys is also important. In labor market interventions where job-related outcomes can take time to materialize, not every benefit is observable immediately after the intervention. The appropriate timeline for follow-up surveys can range from a few months for a job-matching service to up to two years or more for an entrepreneurship program, acknowledging the time it may take for individuals to set up and grow a business. Measuring too early is likely to provide misleading results. In the case of impact evaluations, the timing of the follow-up also has implications on the evaluation methodology.

Another important challenge related to follow-up surveys is locating beneficiaries after they leave the program (or non-participating members of control groups). This is a challenge for many organizations, but it is the only way to determine long-term impacts of a program. Taqueem-supported organizations used a number of strategies for tracking respondents, including gathering accurate contact information at baseline, leveraging social media, providing incentives such as lotteries or cash prizes or using a tracking survey in between the end of the intervention and the follow-up in order to update contact information. Even with these efforts, many organizations faced problems with tracking, which is a particular challenge when working with youth, who tend to change addresses and other contact information with frequency and may be more inclined to migrate than older individuals.

In seeking to find cost-effective and efficient means of collecting data, technology is providing new solutions. Phone-based surveys and messaging, web-based surveys, mobile computers and smartphones, and geographic mapping are increasingly being used for M&E. If used well, these tools can facilitate and often reduce the cost and increase the quality of data collection. In the case of MEDA Maroc, for instance, enumerators carried out impact evaluation surveys using tablets equipped with the Open Data Kit (ODK) survey software (Case Study 6). This mobile data collection technique saved the team valuable time and resources and reduced data errors. Relying on IT solutions also poses challenges, and practitioners need to reserve enough time to fully understand the technology, pilot its use, and train enumerators in its usage before deployment. Moreover, the ordering, programming, editing and connecting of devices can be an arduous and time-consuming task.

Realistic take-up rates

Low take-up and high dropout by beneficiaries can be a major challenge for practitioners. In terms of programming, it reduces the number of participants and thereby may lead to difficulties in meeting program targets. They also pose major challenges for evaluation, particularly impact evaluation. First, having fewer survey respondents reduces the so-called “statistical power” of the evaluation – that is, the probability of detecting an impact if one has occurred. Second, having a high share of dropouts can introduce serious potential biases, because these dropouts are likely to be systematically different from those who stay in the program, but it is hard to know *a priori* how dropouts are different from those who stay in a program. Both the reduced power and the bias can severely undermine an evaluation and minimize the value of its findings. From a programming and evaluation standpoint it is therefore crucial to understand and anticipate take-up and dropout and put in place appropriate mitigation strategies.

Even in a well-implemented program, beneficiaries may be dropping out for reasons external to the intervention itself. In that case, it is extremely useful to understand the reasons for dropout,

which requires locating the respective individuals. For instance, Al-Amal Microfinance Bank found that 87 percent of their dropouts were female, with 45 percent of them indicating that the loan or payment method did not suit them well. AMB can use this information to make adjustments to its products and services accordingly.

Avoiding errors in data collection and recording

Errors in the collection and recording of data will inevitably arise. Errors occur because of interview bias, inadequate methods/questions, processing errors and non-response bias. Once data is collected, steps need to be taken to verify accurateness of replies, this can include a backcheck of a select group of surveys, triangulation and “key judges” for qualitative data.

Data management and analysis

The data collected through the M&E system is only useful if turned into information and conclusions that can lead to action. It therefore needs to be systematically analyzed and interpreted. Managing data is a complex task and needs to be led by a competent M&E officer with a strong background in statistics or econometrics. External specialists with skills in the area can be recruited to complement the project team. Beyond this, there are specific recommendations that come out of our experience with supporting youth-serving organizations that should inform one’s data management and analysis efforts.

First, attention to dataflow is important. The movement of data from one user or system to the next has important implications for the management and quality of the data. In the example of IYF Egypt, data collection started with the program officer at the training site, was then transmitted to a local office where it was entered into a database and finally, was transmitted to global headquarters (Case Study 4). Further complicating matters was the collation and recording of data, which involved several steps from manual data collection on paper-based forms to recording responses in Microsoft Excel, to emailing data to central office, and migrating data to an online cloud-based database. These processes have been complex and lengthy and were a source of delay and poor data quality. In particular, each step required manual inputs which increased the opportunity for data entry errors. Organizations should consider keeping data processes simple and limit the number of people dealing with data.

Second, organizations that collect large amounts of data and that are undertaking evaluations, particularly impact evaluations, need access to more sophisticated data analysis software. Among the supported organizations, a majority used descriptive statistics for analyzing results (see box 6). The tools used for analysis included basic software such as Microsoft Excel to web-based solutions such as customer relationship management software, like Salesforce or GoProve. In the case of more sophisticated analyses, organizations would require access to statistical

BOX X. Refresher on M&E statistics

Descriptive statistics include statistical procedures that can be used to describe the population under study. Frequencies, measures of central tendency (mean and median), and graphs like pie charts and bar charts that describe the data are examples of descriptive statistics.

Inferential statistics is concerned with making predictions or inferences about a population from observations and analyses of a sample. Examples of inferential statistics include regression and correlation analysis.

analysis software such as STATA and SPSS, which allow for inferential statistical analysis. The MEDA Maroc case study used STATA to perform data analysis (Case Study 6).

The introduction of cloud-based M&E systems, while posing some initial adoption costs, promises to help streamline data collection and analysis. Thanks to a cloud-based data management system, IYF has been able to aggregate and disaggregate the data collected in a variety of ways. Program staff can now link beneficiaries to different implementing partners and compare baseline and exit survey results for sets of participants, to better understand which training programs best meet employers' needs. The cloud-based system introduced by LOYAC in Kuwait allowed program staff to enter data from different locations and have "live" access to program data.

Keep in mind that the implementation and ongoing maintenance of any system can be expensive. Many offers of "free" or open source software appear attractive at the outset, but later prove to be more expensive than other options. For example, in the case of IYF's Egypt@Work, Salesforce offered a free license, but IYF had to rely on the expertise of a consultant to configure and adjust the database as the needs of the project changed. Organizations should assess the total costs of any particular software solution, including the costs associated with training staff and hiring consultants with that capacity, before it is adopted.

Sharing and learning

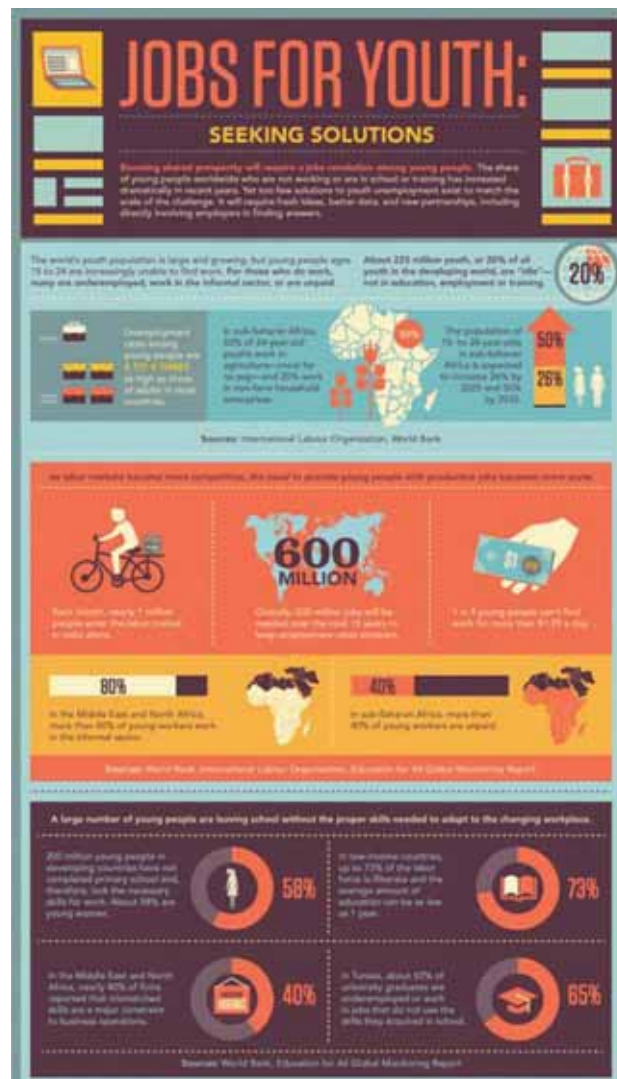
Running evaluations to generate new knowledge is unlikely to have the desired effect if the information and results are not used to their fullest potential. Indeed, there are major challenges in the use of evaluation findings, and many expensive evaluations have failed to lead to significant learning and better programs. In short, evidence is a necessary but not sufficient condition of improved performance and practice (Ramalingam 2011).

Programming and policy-making are complex. Simply presenting information to stakeholders and expecting them to act upon that evidence is unlikely to be successful (Young and Mendizabal 2009). In fact, many practitioners and policy-makers probably tend to be more influenced by their own values and experiences, pressures from supervisors and other groups, and the resources available rather than by evidence. With this in mind, implementing organizations and researchers need to develop a policy outreach strategy to effectively communicate evaluation findings to the broader policy audience.

Planning to share evaluation findings

How findings and recommendations will be communicated should be planned at the same time as the M&E system. This includes planning what audiences will receive the findings, in what form findings will be communicated and in which venues, as well as defining the expected impacts of your sharing activities. In this regard, there are two specific recommendations that come from our experience.

First, it is suggested for M&E teams to develop Policy Influence Plans (PIPs), as commonly requested by the International Initiative for Impact Evaluation (3ie) when funding impact evaluations. For instance, MEDA Maroc developed a PIP in parallel to designing the impact evaluation, identifying individuals in government, donor organizations, and other relevant institutions that would be interested in the findings and that could champion the results of the evaluation to influence future programming on youth employment in the country.

FIGURE 4. Example of infographic

<http://www.un.org/youthenvoy/wp-content/uploads/2013/10/900-YouthJobs-Infographic-English.png>

Second, in considering stakeholder engagement, efforts should begin early. The literature on evaluation use suggests that key stakeholders need to be involved throughout the evaluation process (Patton 2008b). The primary intended users of the evaluation must be clearly identified and engaged at the beginning of the evaluation process to ensure that their interests at the results stage can be identified and addressed. Moreover, evaluators must ensure that these intended uses of the evaluation by the primary audience guide all other decisions that are made about the evaluation process, including defining the evaluation questions.

Tools for communicating findings

Key audiences for evaluation findings include staff and management, partner organizations, government officials, beneficiaries and funding agencies. It is important that messages are adapted depending on the audience, and that messaging is clear, concise and timely. Well-timed communication efforts focusing on “windows of opportunity” – moments when the evaluation results are most relevant to the key stakeholders to make decisions – are therefore

vital for the use of evaluation findings. Al-Amal Microfinance Bank in Yemen, for instance, shares M&E findings in monthly management meetings and makes data accessible to all employees.

When communicating results, a selection of different venues should be used, including presentations for staff and management of the implementing organization, conferences and policy workshops, webinars, websites and blogs, etc. Written reports are still the main medium expected by donors and management. Written reports should include visual displays, not only graphs and charts indicating trends, but also photographs or infographics (see figure 4) which bring information to life. Taqueem-sponsored organizations were also required to supplement M&E reporting with easily digestible case studies or policy briefs, i.e. notes of 3–4 pages summarizing the most important information about the program and its results. For robust evaluation findings from impact evaluations, journal papers remain a key outlet to disseminate findings and raise attention, but their audience is often academic. As such, specifically policy-directed reports should also be developed for such studies.



4. LESSONS ON M&E CAPACITY DEVELOPMENT

In the context of its long-term goal of building evidence of what works for youth employment, one of the key objectives of the Taqeem initiative has been to build the evaluation capacity of organizations working towards employment opportunities for youth in the Arab world. Our capacity-building efforts have focused on working with YSOs to design, implement, and manage M&E systems, as well as using evaluations as a performance management tool. With its expanded approach, Taqeem has aimed to enable organizations to choose the right evaluation tools for their situation and to carry out the M&E needed to improve the effectiveness of their interventions. The following lessons learned from our experience in this area are provided in the hope that more evaluation support initiatives will be developed.

“Monitoring and Evaluation is often times seen as a luxury in projects that is something that you shouldn’t devote a lot of resources to. But what has come across in Taqueem is that it is not a luxury; it is critical to successful program implementation.”

– Clinic participant, Doha, 2011

Capacity for impact evaluation must rest on a solid foundation of more basic M&E capacity

While our ultimate goal was to encourage and facilitate more impact evaluations in order to strengthen the evidence base on youth employment, we recognized early in our approach that few practitioners in the Arab world were ready to carry out this type of robust evaluation. Ignoring smaller organizations with lower capacity in M&E was not a realistic option for Taqueem, since these players represent the vast share of practitioners on the ground with a large potential for learning. This is why we broadened our approach and the content taught in our evaluation clinics, moving from pure impact evaluation content to training on more basic M&E and results-based management. In order to build a strong pipeline for future rigorous evaluations and knowledge generation, it is important first to strengthen practitioners’ foundations in basic M&E.

Success approaches to capacity-building require continuous coaching and mentoring

Training workshops are a good start to introducing practitioners to core concepts behind M&E, but they are not sufficient to induce organizational change. Even if training participants

“Taqeem brings together a mix of people with specific skills when it comes to impact evaluation, monitoring, and evaluation from a variety of institutions that are engaged in this.”

– Technical expert

“It has made me realize the complexity of our program, and I have learnt how to handle the M&E for our programs.”

– Clinic Participant,
Dar es Salaam, 2012

walk away convinced about the importance of M&E, the significance of impact evaluations, and the potential value of stronger M&E and impact evaluation for their organization, it does not mean that they will be able to put this into practice. In addition to technical know-how, many practitioners need support in identifying and prioritizing their M&E needs, matching them with their organizational capacities, choosing the right tools for their organization or program, and supporting the implementation of new M&E practices.

This process requires continuous assistance. This is why Taqeeem assembled a team of technical experts that helped organizations identify their M&E needs, assisted in creating and validating an M&E plan, and provided technical

assistance during the implementation phase of the new M&E strategies (see box 6). Coaches conducted trips to the field and spent an average of 20 days supporting an organization over a period of 10 months. Most support was provided by international consultants, although there was a strong desire by both organizers and supported organizations to engage experts based in the Arab region. Feedback received from supported organizations showed that this coaching and long-term technical advisory support was extremely useful and should have been of longer duration.

BOX 6. Organizations providing technical assistance to Taqeeem members

Swiss Academy for Development
Silatech
JPAL
IPA
Save the Children
Chris Wood Associates
The World Bank
Dalberg Associates

Capacity-building needs to be complemented with financial resources

Even a combination of training and coaching may not be sufficient in making sure that strong M&E systems are put in place. This is most obvious in the case of prospective impact evaluations, which are often expensive due to the associated data collection and consulting needs. Importantly, project budgets often do not foresee the sufficient expenditures for data collection and other M&E-related expenses. Through Taqueem, we provided organizations within the Taqueem CoP with grants to cover direct costs of data collection, in addition to the in-kind support of paying for technical experts. In addition, the ILO provided a pool of funding to organizations ready to carry out a robust evaluation, which included a competitive “Fund for Evaluation in Youth Employment” which provided grants of US\$5,000–180,000.

M&E training should be provided to more than one person in an organization

Offering regional training workshops has the inherent challenge that only selected representatives of an organization can often be reached, especially when targeted organizations are coming from different countries. This poses a potential challenge with ensuring that learning is institutionalized, rather than remaining with the individual directly receiving training. In its initial trainings, Taqueem usually managed to have at least two representatives from an organization attend, one at the senior management level and one at the technical M&E level. Nonetheless, Taqueem faced a repeated issue with M&E officers leaving their jobs and the organization after receiving training and coaching. Among our supported organizations, Education for Employment, LOYAC, and the International Youth Foundation all faced this challenge.

Such departures can leave behind a vacuum and render the implementation of new M&E approaches more difficult, and underscore the importance of not investing all M&E expertise in just one person. In particular, senior management need to be actively involved in the development and implementation of new M&E activities, so that the M&E knowledge and process are embedded within the organization.

“Meeting with other NGOs and other institutions from the Arab world helped us compare what we are doing. That’s a mutual exchange that’s very positive.”

– NGO participant

“I think the networking part is very important to emphasize. People are leaving here and they form friendships, they really have a better sense what other people are doing, what other countries’ needs are, and a better overall sense of what they want to do with their own work.”

– Technical expert

Creating networks of practitioners fosters exchanges on evaluation and programming

A key feature of the Taqueem Initiative’s capacity-building efforts was the initiation of a practitioner network. The Community of Practice provided a forum for sharing needs and experiences on

M&E, but also on youth employment programming more broadly, allowing for peer-to-peer learning in addition to the services provided directly by Taqueem.

The peer-learning and network effects have been further facilitated by Taqueem's sector-specific approach to M&E; that is, activities targeted at practitioners in the youth employment field. While it is true that M&E and impact evaluation are equally important across all sectors, the fact that all organizations and experts involved had the same background and thematic focus enhanced group cohesion and practitioners' ability to relate to the concepts, M&E approaches, and lessons learned.

Leveraging academic and research partners is important

While the fundamentals of M&E are relatively easy to convey to organizations through training, they will face technical issues upon implementation that require particular expertise and experience with M&E. This is particularly true in regard to impact evaluation. When planning an impact evaluation, issues of counterfactual selection, quality data collection, and validity of results are generally not possible for implementing organizations to handle in-house without external assistance. Research partnerships can bring much needed evaluation skills and expertise to resolve these constraints. They can also be essential for establishing an impact evaluation's credibility among key stakeholders. Research partners also help make connections that allow the organization to more widely disseminate evaluation results.

When making the choice of a research partner, whether an individual or an organization, it is important to consider a number of factors:

- Pay attention to values, not just technical expertise. It is important that local communities are involved in the research, not just treated like test subjects.
- The type of research partner you are working with and their research interests may affect the impact evaluation. Often, researchers have their own research agendas, determined largely by research priorities in their given field. It is important that the research interests of the research partner fit the learning objectives of the implementing organization.
- Be open to a variety of potential research partners, including for-profit research firms, non-profit research organizations, universities, or colleges.
- Consider establishing a relationship with a research partner early, so programming and research can go hand-in-hand.

The Taqueem Initiative's administrative team played the role of providing links to research partners, as the choice of which partner to work with was a necessary but often difficult one to make for implementing agencies. Taqueem's matchmaking role produced a number of fruitful partnerships between organizations and research institutions, including Innovation for Poverty Action (IPA), the Abdul Latif Jameel Poverty Action Lab (J-PAL), Graduate School Institute Geneva, Georgetown University and the American University of Cairo. Another good example of bringing researchers and practitioners together is IPA's matchmaking program (see box 7).

BOX 7. IPA's matchmaking program

Innovations for Poverty Action's matchmaking programs facilitate conversations and collaborations between researchers and practitioners, with seed-funding awarded to the most promising research partnerships. Practitioners seeking to partner with researchers to develop and evaluate innovative programs and researchers looking for partners interested in piloting and studying research ideas can apply to this matchmaking program. The objective of the matchmaking is to provide an arena for the development of new research partnerships and to support rigorous testing of innovations designed to be scalable and policy-relevant.



5. MOVING FORWARD: STRENGTHENING M&E OF YOUTH EMPLOYMENT PROGRAMS

.....

Drawing on the Taqeem experience outlined in this volume, some general recommendations emerge on how to further strengthen learning and knowledge generation in the youth employment field. The following recommendations are intended not only for partners in the Taqeem Initiative, but for the youth employment community more broadly. They are particularly relevant for development partners and other agencies interested in strengthening knowledge and results of youth employment interventions.

Expand investments in M&E capacity-building. Demand for assistance in improving capacity in monitoring, evaluation and impact evaluation and strengthening results-measurement frameworks of youth employment interventions remains largely unmet. Between 2010 and 2013, the ILO received over 500 requests for assistance from almost 50 countries across Africa and the Arab world. Meeting this persistent demand and broadening the level of M&E services provided will require partnerships at the international and local level.

In order to meet the widespread demand, and given that effective M&E capacity-building requires continuous assistance, future capacity-building activities should leverage local M&E service providers. The Taqeem Initiative is planning to establish sub-regional hubs that could host M&E training initiatives with a focus on youth employment, identify and maintain a database of local M&E experts, and provide targeted M&E technical assistance to local organizations. Similar efforts by other institutions can serve as an example, such as the CLEAR initiative (Regional Centers for Learning on Evaluation And Results) aimed at strengthening partner countries' capacities and systems in monitoring, evaluation, and performance management. A stronger engagement of national and regional evaluation institutions such as the MENA Evaluators Network (EvalMENA) and the Interagency Regional Evaluation Network for Arab States, able to provide local and on the ground support can further enhance long-term sustainability. The Taqeem experience suggests that M&E assistance should have a thematic lens (i.e., on youth employment) to facilitate peer-learning among supported organizations; thus, existing and potential service providers may need to be supported to deepen their understanding of youth employment to make targeted M&E services part of their portfolio.

M&E capacity-building and targeted funding should be regular and predictable. In addition to the ILO, there have also been efforts by other institutions to provide capacity-building and financial assistance for impact evaluations of youth employment programs, including by the World Bank, J-PAL, 3ie, and others. However, one key challenge so far among all the existing support mechanisms is their lack of regularity and predictability. Usually, capacity-building and funding

opportunities are announced ad hoc and with relatively short notice, making it often difficult for practitioners to take advantage of these opportunities. For capacity-building and funding support on M&E and impact evaluations to be most effective, practitioners need to know what and when they will be available, so they can take it into consideration for their own planning purposes. Future evaluation support efforts should therefore adopt a multi-year framework.

Establish networks of local researchers or research institutions with capacity to support impact evaluations and related capacity development. A key challenge for many evaluation efforts remains the scarcity of qualified local professionals in some countries, and even where those exist they are not necessarily known by practitioners. Future capacity-building efforts should continue to explore ways to involve local researchers and research institutions and actively link them to programs on the ground. Concretely speaking, scholars in local research institutions and universities could be paired with beneficiary organizations showing an appropriate level of maturity and interest to conduct an impact evaluation in order to help develop a concept note that can be used to seek funding.

Donors and funders should encourage implementing agencies to include M&E capacity-building in program budgets. In order to effectively carry out M&E as key component of every program, it is naturally crucial that a portion of the program budget be reserved for M&E staff, data collection, and knowledge-sharing. Recognizing that local M&E expertise is still limited in many countries, and that organizations' M&E needs are diverse and dynamic, funders should promote that a greater portion of the M&E budget be used for capacity-building and external technical assistance. This would give organizations more leeway to build strong M&E systems and evaluation designs, thereby enhancing program quality over the medium term.

Continue to invest in impact evaluations. Since the first analysis of the Youth Employment Inventory was carried out in 2007, much has happened to generate new knowledge on youth employment. The scarcity of evidence on youth employment programs in developing countries, and in particular in Africa and the Middle East, has sparked considerable efforts to invest in robust evaluation. In addition to the ILO, the World Bank, 3ie, J-PAL, IPA, DFID and others have supported a series of new evaluations whose results have started to emerge. The Youth Employment Inventory itself has grown to over 700 interventions captured in the database (see box 8) and a series of review papers and meta-analyses have emerged summarizing recent evidence on youth employment interventions (see, e.g., Cunningham, Sanchez-Puerta, and Wuermlli 2010; Ibarrarán and Rosas 2009; Cho and Honorati 2012; J-PAL 2013; UNF 2013; USAID 2013a, 2013b).

Despite this progress, many knowledge gaps remain. The Arab region still lags behind in terms of quality evaluations on youth employment, despite encouraging examples from Jordan, Tunisia, and Morocco (Groh et al. 2012; Premand et al. 2012; Crépon et al. 2011). Overall, much more knowledge is needed on non-training-related interventions, the relative effectiveness of different design features (e.g. dosage, delivery channel, pedagogies), the role of context (e.g. urban vs. rural, informal vs. formal economy), and the effect of employment programs on non-labor market outcomes.

Growing the evidence base on youth employment should rely on predictable funding for impact evaluations. As for M&E-related capacity-building, one key barrier for practitioners is the lack of predictable funding for impact evaluations. For example, even if a sudden funding opportunity arises, an NGO currently designing a new program may not be ready to submit an impact evaluation proposal. Without knowing that there will be another possibility to apply for funding 6 or 12 months later, this situation may provide a lower incentive to plan for a prospective

BOX 8. The Youth Employment Inventory (YEI)

The YEI is a dynamic database that provides comparative information on youth employment interventions worldwide. Comprising more than 700 youth employment programs from around 90 countries (growing), YEI documents program design, implementation, and achieved results. The programs included range from interventions for improving the labor market information base (counseling, job-search skills) to programs that aim at increasing the demand for youth labor (wage subsidies and public works) and those focusing on improving chances for young entrepreneurs (providing financial, technical, and training assistance).

Originally initiated by the World Bank, the YEI today is a joint effort of the German Ministry of Economic Cooperation and Development, the Inter-American Development Bank, the International Labour Organization, World Bank, and Youth Employment Network.

Organizations can submit their programs to the database at: <http://www.youth-employment-inventory.org>

evaluation. Many of the recent rigorous impact evaluation studies have been opportunistic, made possible in the context of a particular funding opportunity. A more systematic and long-term approach would allow for expanding the numbers and geographical scope of credible evaluations, and could enable the targeting of specific knowledge gaps over time.

A more predictable model for impact evaluation funding in the area of youth employment could be inspired by USAID's Development Innovation Ventures (DIV). DIV uses a three-tiered staged finance model, ranging from i) seed-financing to establish proof of concept, ii) evaluating solutions at scale, and iii) testing program effectiveness during its transition to widespread adoption. Since applicants can apply to any stage and to higher stages once success can be demonstrated at lower levels, the model provides predictability and funding for evaluation in line with the dynamics of a growing intervention. Providing such flexibility would require funders of evaluations to adopt a longer time horizon, so that there can be repeated calls for proposals or rolling basis applications over several years.

Impact evaluations on youth employment should be systematically complemented by cost-benefit analyses. While the evidence base on youth employment has been growing, there remains a serious scarcity of knowledge on program costs. This lack of information impedes a more detailed perspective on the returns of youth employment interventions in terms of cost-benefit comparisons despite the fact that this information is extremely relevant from a policy perspective. Funding frameworks for impact evaluation studies should also provide incentives to build in robust cost-effectiveness and cost-benefit analyses.

Strengthen structured iterative learning approaches.

Those operating in the youth development sphere tend to focus on questions of program design in discussing strategies to enhance youth employment (and development interventions more broadly), and, in doing so, sometimes neglect the crucial importance of quality implementation to build successful programs. Especially when a program is unsuccessful, it is crucial to differentiate between whether this

“

Without impact evaluation, we don't know what works. We don't know where we failed, we don't know where we have succeeded.”

– NGO participant from Taqueem evaluation clinic

was due to bad design or simply bad implementation. This is why many evaluation practitioners favor mixed-method approaches, so that evaluations do not only answer whether a program was successful (or not), but also grasp the reasons behind this success or failure.

The conversation about the importance of implementation has recently gained momentum. World Bank President Jim Kim stated that helping advance a “science of delivery” is the next frontier for the World Bank Group, recognizing that most failures happen at delivery. He stressed that effective delivery demands context-specific knowledge, constant adjustments, and a relentless focus on the details of implementation, a message echoed by others now emphasizing the importance of “delivery” (Kim 2012; see also Barber, Kihn, and Moffit 2011; McKinsey & Company 2013).

Against this background, traditional M&E and impact evaluation – despite being crucial for learning – may not suffice. By nature, these approaches rely on a fairly static intervention to assess whether implementation is going according to plan, whether targets have been achieved, and whether causal impacts can be seen. Yet, practitioners often stress the importance of context-specific programming, flexibility, and the ability to adapt to changing circumstances, thus underlining the critical role of dynamic learning. Rather than limiting M&E to assess the effectiveness of *ex ante* project assessment, learning could be strengthened by fostering structured feedback loops during implementation. Such an approach would be particularly relevant to youth employment interventions, where many knowledge gaps on adequate design features and implementation arrangements remain.

To this end, we need to experiment with different design and implementation options, monitor outcomes, and adapt accordingly. To overcome the limitations of traditional M&E and impact evaluations, Pritchett et al. (2013) have proposed the concept of “structured

experiential learning.” The core idea of structured experiential learning is to build learning objectives into the cycle of project design, implementation, and evaluation. Concretely, this means articulating project or program alternatives, simultaneously trying out these design alternatives, and then adapting the project sequentially based on the results. In practice, this could mean that, rather than making an *ex ante* decision as to whether to provide cash- or in-kind grants as part of an entrepreneurship project, the project would provide cash to some

“The real challenge is rarely a lack of new ideas and innovative concepts. Nor is it the availability of funds. Rather, it’s the difficulty of implementing simple, pragmatic, scalable solutions in a sustainable way.”

– Norbert Dörr,
McKinsey & Company

(Source: McKinsey & Company 2013)

beneficiaries and in-kind grants to others. Similarly, it could mean creating different groups of beneficiaries who receive different amounts of training, to see what how many hours of training work best. As such, within-project variations in design or implementation can serve as their own counterfactual, with the potential to reduce the cost of evaluation and increase the direct usefulness of evaluation to implementing agencies. By experimenting with different program features and monitoring them closely, dynamic feedback loops can be created to help continuously adapt and improve an intervention. Because the learning objectives

were built in from the start, the information created in this process is directly useful to the implementing agency.

Taqeem and future work to strengthen M&E in the Arab World

The recent investment by the ILO and its partners in impact evaluation and M&E more broadly is leading to an improvement of the evidence base for effective design and implementation of youth employment programs. The long-term objective is that the findings and recommendations coming out of this work will lead to the creation of evidence-based policy for youth and that these policies will transform into scalable and sustainable employment impacts. With this in mind, the future work coming out of the Taqeem Initiative on this topic will revolve around three strategies to achieve policy influence:

1. **Anchor the capacity-building and technical assistance model in local evaluation institutions.** The objective is for these institutions to become M&E and impact evaluation “hubs” for youth-serving organizations in their regional community. Evaluation clinics and other technical assistance services will help foster a “critical mass” of M&E and impact evaluation expertise and support establishing stronger M&E systems of influential government and NGO programs.
2. **Continued investment in knowledge generation through high-quality impact evaluations.** Taqeem will continue to support practitioners in the design and implementation of robust evaluations in order to fill remaining knowledge gaps regarding which types of interventions work, which do not, and why.
3. **Policy messages to key stakeholders.** As the knowledge base improves, efforts will be made to ensure emerging evidence is translated into policy. Key officials within influential institutions and governments will be targeted to foster the dialogue on what works and what does not. Evaluation findings will be reinforced through meta-analysis and synthesis of results of youth employment interventions globally.

In order for this effort to be successful, a support coalition of technical specialists, donors, researchers, and academics from leading evaluation initiatives will be built. These partners will collaborate through the Taqeem Evaluation Council, together working to build capacity of local institutions, providing guidance methodology, and assessing outcomes in way that builds the evidence base and provides concrete recommendations about what works on youth employment.

REFERENCES

Adato, M. 2011. *Combining quantitative and qualitative methods for program monitoring and evaluation: Why are mixed-method designs best?* World Bank, Washington, DC. <https://openknowledge.worldbank.org/handle/10986/11063>

Bamberger, M., Rao, V., and Woolcock, M. 2010. *Using mixed methods in monitoring and evaluation: experiences from international development*. Policy Research Working Paper No. 5245 (Washington, DC: The World Bank). Available from: http://www-wds.worldbank.org/servlet/WDSCContentServer/WDSP/IB/2010/03/23/000158349_20100323100628/Rendered/PDF/WPS5245.pdf

Barber, M., Kihn, P., and Moffit, A. 2011. *Deliverology: From idea to implementation* (McKinsey and Company).

Betcherman, G. et al. 2007. *A review of interventions to support young workers: Findings of the youth employment inventory*. SP Discussion Paper No. 0715 (Washington, DC: The World Bank). Available from: <http://www.youth-employment-inventory.org/downloads/1.pdf>

Cho, Y. and Honorati, M. (2012). *Entrepreneurship programs in developing countries: A meta-regression analysis* (Washington, DC: World Bank).

Crépon, B. et al. 2011. *Impact of microcredit in rural areas of Morocco: Evidence from a randomized evaluation*. Working Paper. American Economic Journal: Applied Economics 2015, 7(1): 123–150 <http://dx.doi.org/10.1257/app.20130535>

Cunningham, W., Sanchez-Puerta, M. L., and Wuerml, A. 2010. *Active labor market policies for youth: A framework to guide youth employment interventions*. Employment Policy Primer No. 16. (Washington, DC: World Bank). Available from: http://siteresources.worldbank.org/INTLM/214578-1103128720951/22795057/EPPNoteNo16_Eng.pdf

Groh, M. et al. 2012. *Soft skills or hard cash? The impact of training and wage subsidy programs on female youth employment in Jordan*. Impact Evaluation Series No. 62 (Washington, DC: World Bank).

Hempel, K. and Fiala N. 2011. *Measuring success of youth livelihood interventions: A practical guide to monitoring and evaluation* (Washington, DC: Global Partnership for Youth Employment). Available from: <http://www.gpye.org/measuring-success-youth-livelihood-interventions>

Ibarrarán, P. and Rosas Shady, D. 2009. "Evaluating the impact of job training programmes in Latin America: Evidence from IDB-funded operations." *Journal of Development Effectiveness* 1(2): 195–216. Available from: http://www.iza.org/conference_files/ELMPDC2009/ibarraran_p4263.pdf

ILO 2012. The youth employment crisis: A call to action, *Resolution and Conclusions of the 101st Session of the International Labour Conference* (Geneva: International Labour Organization).

ILO 2013. *Global employment trends for youth 2013* (Geneva: International Labour Organization).

J-PAL 2013. J-PAL Youth initiative review paper.

Kim, J. 2012. *Delivering on development: Harnessing knowledge to build prosperity and end poverty*. <http://www.worldbank.org/en/news/speech/2012/10/08/delivering-development-harnessing-knowledge-build-prosperity-end-poverty>

Leeuw, F., and Vaessen, J. 2009. *Impact evaluations and development: NONIE guidance on impact evaluation*. Washington, DC: Network of Networks on Impact Evaluation. Available from: http://siteresources.worldbank.org/EXT/OED/Resources/nonie_guidance.pdf

McKinsey & Company 2013. *The art and science of delivery. Voices on society*.

Patton, M.Q., 2008a. State of the art in measuring development assistance. Address to the World Bank Independent Evaluation Group Conference, 10 April, Washington, DC. Available from: <http://www.worldbank.org/ieg/conference/results/patton.pdf>

Patton, M.Q., 2008b. *Utilization focused evaluation*. 4th ed. (Thousand Oaks, CA: Sage Publications).

Premand, P. et al 2012. *Entrepreneurship training and self-employment among university graduates: Evidence from a randomized trial in Tunisia*. World Bank, Policy Research Working Paper 6285

Pritchett et al. 2013. *It's all about MeE: Using structured experiential learning ("e") to crawl the design space*. Center for Global Development, Working Paper 322.

Ramalingam, B. 2011. *Learning how to learn: Eight lessons for impact evaluations that make a difference*. Overseas Development Institute, Background Note.

USAID 2013a. *State of the Field Report: Examining the evidence in youth workforce development*.

USAID 2013b. *State of the Field Report: Holistic, cross-sectoral youth development*.

United Nations Foundation (2013). *A roadmap for promoting women's economic empowerment*.

Young, J. and Mendizabal, E. 2009. *Helping researchers become policy entrepreneurs. How to develop engagement strategies for evidence-based policy-making*. Overseas Development Institute. Briefing Paper 53.



PART II

CASE STUDIES



CASE STUDY 1

AL-AMAL MICROFINANCE BANK, YEMEN



Complementing a microfinance bank's management information system (MIS) with targeted M&E to better understand the social impact of lending on the Bank's clientele

KEY MESSAGES

- In designing an M&E system, be mindful of the organizational capacity to implement it. Prioritize information needs and build basic data collection tools and M&E system around priorities before expanding.
- Secure buy-in from both senior management and front-line staff to ensure successful implementation of a new M&E system.
- Technology can be a powerful tool in M&E, allowing organizations to streamline processes, ease the workload on collecting the data, and minimize the number of data entry errors.

Background

Al-Amal Microfinance Bank (AMB) seeks to reduce poverty and unemployment in Yemen by providing a range of financial services, including credit, savings, and insurance, to Yemen's unbanked population. The Bank has 18 branches throughout the country, spread over nine governorates. As of April 2013, AMB had served over 80,000 clients and disbursed over 67,000 loans totalling US\$18.1 million. Women constitute over 60 percent of the Bank's borrowers.

In 2009, AMB and Silatech established the Youth Loan Fund, with an initial capital of US\$2 million, to distribute micro loans and business development services to young Yemenis aged between 18 and 30. The first youth-focused loan program in Yemen, the Youth Loan Fund was developed to support youth through the transition from micro enterprise to small business, with a strong emphasis on long-term sustainability and job creation. Typically, loan terms last between six and twenty-four months, with loans ranging from US\$50 to US\$1,000 per applicant.

Taqeem support

AMB joined Taqeeem to build a body of evidence on the effectiveness and success of its Youth Loan Fund. While the Bank had put in place the basic building blocks of an M&E system before joining, it identified four main areas for improvement:

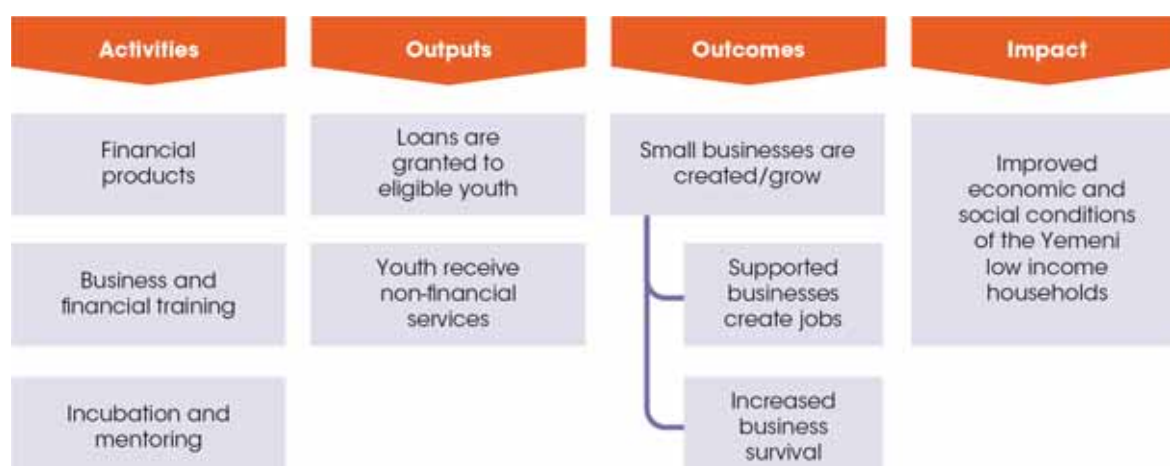
1. **Expanding data collection beyond loan monitoring:** AMB's data collection had focused on financial measures such as loan volume, number of active borrowers, and repayment rate. However, it only collected a few socio-economic indicators focused on measuring well-being and economic status.
2. **Strengthening the management information system (MIS):** The Bank's MIS system required significant changes to adapt to a revised M&E system. Using technology to collect borrowers' financial data as well as social outcome data could further streamline their processes and drive efficiency.
3. **Enhancing data collection capabilities:** AMB relied on its loan officers to collect and record data through loan application forms, home visits, and regular contact with youth lenders. However, loan officers had little experience with qualitative data collection, conducting more sophisticated household surveys, or using mobile data collection devices.
4. **Measuring impact:** AMB had no reliable information regarding the real impact of its microfinance services on medium- and longer-term outcomes such as job creation or business survival.

M&E system

Results chain

By providing access to financial and non-financial services, the Youth Loan Fund program aims to support the creation and growth of small businesses, thereby increasing income and job opportunities and ultimately reducing poverty in Yemen (figure 1).

FIGURE 1. Overview of the results chain for AMB's Youth Loan Fund



Data collection instruments

The basis for AMB to measure its achievements is a set of well-defined indicators linked to each output and outcome (table 1). Many of the data collection instruments used for output monitoring were already in place at AMB when they joined Taqueem, but tools for measuring outcomes were largely absent. By mid-2013, a number of data collection instruments had been piloted and were being rolled out.

TABLE 1. Description of data collection tools and indicators for the Youth Loan Fund

Instrument	Main indicators measured by the instrument	Description
Application form (or baseline youth survey)	<ul style="list-style-type: none"> Percentage of youth clients benefiting from business development services Number of youth clients aged 18–30 Average loan size Demographic information: gender, rural/urban, household income, education level 	Self-administered paper form filled out at the loan application stage that includes information on demographics, family situation, business and loan history, credit history, current income, and budget and assets. Data are subsequently manually entered into the MIS
Follow-up youth survey	<ul style="list-style-type: none"> Average enterprise expenditure per week Value of enterprise New enterprises operating after two years Household health expenditures over last year Number of start-ups created Number of jobs created 	A follow-up survey repeating questions from the baseline survey but adding questions on employment and health. The survey is issued to a sample of borrowers only. The follow-up survey is administered at the loan completion, as well as two years after the loan completion. Surveys are conducted on tablet devices and answers are uploaded automatically to the MIS
Progress out of Poverty Index (PPI)	<ul style="list-style-type: none"> Number and percentage of youth living in poverty Percentage of households sending children to school 	A poverty-monitoring tool developed by the Grameen Foundation and tailored for Yemen. Ten questions are used to determine a score. Administered at the time of applying for a loan and at loan repayment (additional follow-up one year after final repayment planned to measure change in poverty status). Administered as a paper form completed by the Loan officers with the client, and entered manually into the MIS
Client satisfaction survey	<ul style="list-style-type: none"> Percentage of young clients satisfied with AMB services during the loan period 	A survey to measure how satisfied young clients are with AMB services and administered once with each client. The survey is planned to be implemented on tablet devices, allowing loan officers to collect data when visiting clients and to send collected data automatically to the MIS database

(continued)

TABLE 1. (continued)

Instrument	Main indicators measured by the instrument	Description
Focus groups	<ul style="list-style-type: none"> ■ Motivation and confidence in the client groups and anecdotal evidence of impact in the wider society 	Group meetings of available clients in the same location who have paid off their loans to discuss personal stories of the impact of the program. Groups are arranged and facilitated by the M&E officer

Data management

AMB had constructed a comprehensive management information system which tracks and records all financial transactions and assists in project management. This database became the backbone of the M&E system as a repository for storing, analyzing, and distributing results and findings. To ensure this capability, the database had to undergo an upgrade to include socio-economic fields and add the capacity to track longitudinal information.

Loan officers are responsible for collecting data from borrowers through loan application forms, home visits and other communication channels (figure 2). AMB's research and development department is responsible for entering all data into the MIS database and analysing findings. They are also responsible for coordinating focus groups. Follow-up surveys are expected to be carried out by an external data collection firm.

In the future, loan officers are expected to collect information through tablet devices allowing data to be directly synchronized to the database. This digital data capture will ease the workload on collecting the data and will minimise the number of data entry errors.

M&E results

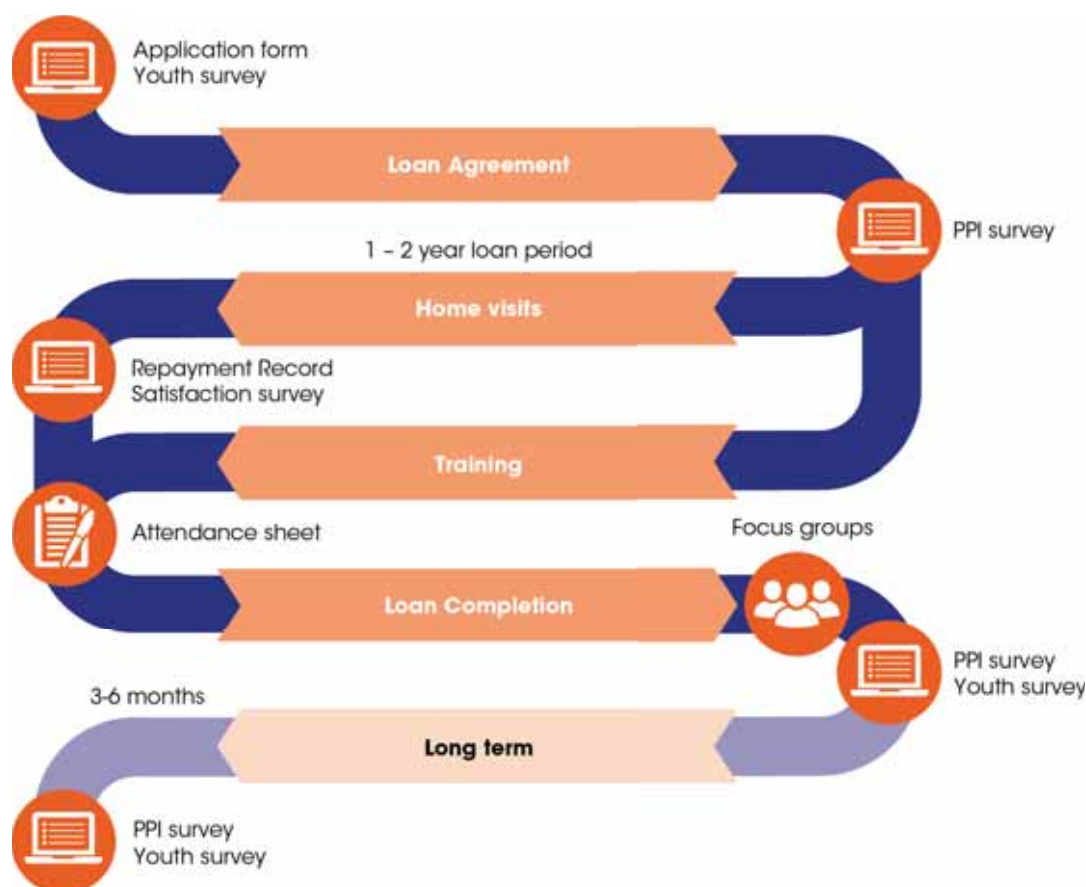
An initial analysis of the existing data yielded the following findings:

- **Beneficiary characteristics:** The Youth Loan Fund targets young people aged between 18 and 30. The majority of youth clients are women (62 percent) and the average age is 25.6 years.
- **Dropouts:** For various reasons, youth clients drop out of the program after they receive the loan. Overall, the dropout rate is 24 percent; of these, 87 percent are women. Nearly 45 percent of dropout clients justify dropping out because of the type of loan or the payment methods. These dropouts may be as a result of the wrong type of loan being agreed to in the first place and clients only realizing the implications afterwards. Moving forward, it will be crucial for AMB to understand why young people are withdrawing from the program.
- **Poverty index:** AMB's initial PPI data also suggest that almost 30 percent of women for which data are available ($N = 268$) have seen a negative change occur in terms of their poverty situation during the loan period. For men, this percentage reaches 35 percent ($N = 48$). While this could be due to a number of factors, such as social and political instability,



“ The first youth-focused loan program in Yemen, the Youth Loan Fund was developed to support youth through the transition from micro enterprise to small business, with a strong emphasis on long-term sustainability and job creation

FIGURE 2. Overview of the data collection workflow of AMB's M&E system



a deeper analysis would help AMB better understand these results. For instance, by looking into the type of loans given to youth clients for whom a negative change has been observed, the Bank could better align loan product characteristics with client needs. Early results also show that relatively more women are experiencing a positive change than men (50 vs. 38 percent) in terms of their poverty situation, indicating that gender dynamics may influence the effectiveness of the youth loans in Yemen.

Lessons learned

Keep it simple

In designing an M&E system, it is important to ensure that it is aligned with the capacity of the organization to maintain it. A limited number of simple data collection tools and indicators are normally enough to start an effective M&E system. Once the system is running, other tools and indicators can be added. AMB's original M&E plan included over 50 indicators and 10 data collection instruments. In the end, AMB cut its indicators down to 20 and discarded plans for video focus groups and an exit survey. Data collection tools such as the loan application form have been adjusted to allow for higher-quality data to be collected from already existing tools and technological solutions such as the use of tablets are being introduced to reduce time and error.

“Early results also show that relatively more women are experiencing a positive change than men in terms of their poverty situation, indicating that gender dynamics may influence the effectiveness of the youth loans in Yemen

Put in place follow-up procedures

Following up with beneficiaries after their loan has been repaid is the only way to determine if medium- or long-term outcomes have been achieved. However, this proved challenging for AMB as it faced many competing business priorities. Moreover, beneficiaries can be hard to track in Yemen as they often live in hard-to-reach rural communities and are highly mobile. Options to make follow-up easier include:

1. Constantly update contact databases. To save time and expand reach, such tracing exercises can be conducted using bulk SMS messaging.
2. Offer a small incentive (e.g., free phone credit) to youth who fill out a survey.
3. Hire a professional data collection firm which can help you trace contacts and provide advice on survey design and methodology, data collection and analysis.

Ensure buy-in from implementing organization

In AMB's case, the call for M&E reform was a top-down initiative, and securing buy-in from senior management was key before any changes could take place. Part of that buy-in was incentivized by the opportunity for funding and external partnerships. However, AMB's ability to absorb M&E changes was largely due to the personal commitment of the Bank's CEO, who himself attended Taqueem Clinics and pushed for changes to the Bank's approach to M&E.

On the other hand, the introduction of M&E into AMB created a new set of processes, resource demands and competing interests, increasing resistance among many front-line employees at AMB. Loan officers are at the front line of data collection, which increased their workload. The newly formed research and development unit is responsible for the M&E system, but they face a lack of human and financial resources needed to accomplish the ambitious plans. A consultative process across departments has been initiated within AMB to discuss the implementation of new M&E tools and help increase ownership among staff. Continuous capacity building and support for staff will be needed to anchor M&E within the organization. AMB also plans to highlight the achievements of its departments and thereby increase the perceived value of supporting M&E across the organization.

Next steps

Impact evaluation

The Taqueem initiative has allowed AMB to build a stronger M&E system while also learning about impact evaluation. In turn, AMB, in partnership with Georgetown University and Silatech, initiated

an impact evaluation of the Youth Loan program in late 2013. The study seeks to assess the impact of youth-oriented microcredit on poverty, business growth, market behaviour, and financial risk of participants. The evaluation uses an RCT approach to examine the impact of participation in the program on (i) employment and economic opportunity; (ii) income and consumption; and (iii) vulnerabilities. The impact evaluation is financed by the Qatar National Research Foundation.

Database development

AMB's MIS system will receive an upgrade to further refine the database to accommodate M&E survey data collection. The design of the follow-up survey is being finalized, as are the questionnaires for the satisfaction surveys. Loan officers will also start to pilot the use of tablets to collect data during the visits with the client. Data input will be synchronized with the database.

Using the results

Given its interest in scaling-up and increasing the capital of the Fund, AMB intends to further enhance the use of its findings from M&E activities. This includes stronger outreach to better communicate the project's outcomes through targeted reporting to donors, presentations of results in public events, and highlighting findings in the Bank's newsletters, annual reports, and website. Moreover, AMB seeks to improve program delivery by assessing the performance of each area of the process. To this end, findings will be shared and discussed in monthly senior management meetings, and all AMB employees will have access to the database in order to strengthen the awareness of program performance.

Continued capacity-building

AMB's research and development team have participated in three Taqueem Evaluation Clinics where they have learned basic theory on M&E and impact evaluation. In the future, AMB will continue to build the M&E capacity of its research and development unit including by participating in trainings on impact evaluation.



CASE STUDY 2

EDUCATION FOR EMPLOYMENT FOUNDATION, MOROCCO



Revamping an M&E system to meet the needs of a growing NGO in becoming more results-driven

KEY MESSAGES

- Conducting a follow-up survey amongst program beneficiaries is strongly recommended to track longer-term outcomes.
- When following up with a large number of beneficiaries, sampling might be advisable, as opposed to tracking all beneficiaries. Sampling should be done by random selection to minimize potential bias in results.
- While web-based surveying can reduce time and resource requirements, it also poses numerous challenges, such as low response rates and potential response bias.

Background

Education for Employment (EFE) is a network of local non-profits in Egypt, Jordan, Morocco, Palestine, Tunisia, and Yemen, which are supported by organizations in the United States and Europe. The EFE network addresses the youth employment challenge in the Middle East and North Africa (MENA) region, seeking to create job opportunities for unemployed youth by providing professional and technical training that leads directly to career-building jobs.

EFE Maroc was founded in 2007 and currently runs four distinct training programs, one of which is called Finding a Job Is a Job (FJIJ). FJIJ is a three-day employability training targeting unemployed youth, often youth who are transitioning from school to work. FJIJ teaches job-search skills, how to improve CVs and cover letters, and how to perform in a job interview. Every course brings together one trainer and 20–25 participants. EFE Maroc's goal is to train 12,500 youth in FJIJ over a four-year period, which started in 2011. EFE Maroc partners with youth centers and universities in the delivery of the project, and training is delivered in six regions across Morocco. FJIJ training takes place under the umbrella of the Al-Morad Initiative, a wider project financed by the MasterCard Foundation.

Taqeem support

The monitoring and evaluation (M&E) function at EFE falls under the Performance Management department, which is a centralized function overseen by EFE's headquarters in the United States. Each EFE affiliate has an M&E manager responsible for all projects within the country. At the time of joining the Taqeeem Initiative, EFE Maroc relied on complete M&E systems for all but their newest training program, FJIJ. EFE Maroc therefore joined Taqeeem, seeking to develop a comprehensive and innovative monitoring and evaluation plan that would provide a harmonized approach to collecting and analyzing outcome data for FJIJ.

M&E system

Results chain

The FJIJ program is based on the assumption that the education, knowledge, and skills provided by the public education system in Morocco are often insufficient for enabling graduates to find quality employment. It also assumes that jobs are available for young people that have the right skills. With this in mind, EFE seeks to meet two outcomes: (1) young people gain skills and knowledge on how to conduct an effective job search; and (2) the job search is successful, leading to decent employment (figure 1). The long-term impact of the project is to contribute to socio-economic inclusion of underprivileged youth in Moroccan society.

FIGURE 1. Overview of the results chain for EFE Maroc's Finding a Job Is a Job (FJIJ) project



Data collection instruments

EFE Maroc rolled out a series of revised data collection tools between 2012 and 2013 to better capture the outcomes of the program. The M&E plan for EFE's FJIJ project includes five instruments which measure a series of indicators, as described in table 1 (see also figure 2).

Data management

EFE Maroc's M&E database is a Microsoft Excel spreadsheet in which all administrative and survey data are input, cleaned, and analyzed. While not an ideal tool, EFE Maroc decided to use Excel for this operation as it provided a simple, low-cost solution and one that staff already had the skills to operate.

TABLE 1. Description of data collection tools and main indicators of the M&E system for FJIJ

Instrument	Main indicators measured by the instrument	Description
Application form and interview (baseline)	<ul style="list-style-type: none"> Number of FJIJ applicants Level of job-search knowledge and skills Attitude towards employment Motivation to get a job 	An online application form asks around 15 questions on socio-demographics, work status, education, and motivation, based on which EFE determines eligible applicants and contacts them for a short interview. Data are exported to Excel and filtered
Attendance sheet	<ul style="list-style-type: none"> Number of youth participants enrolled Number of youth who graduate from program Number of trainings delivered 	Attendance sheet administered every session by the trainers. Important since graduation certificate is given only to those attending at least 60% of the sessions. After every training, trainers send a scan of the attendance sheet to the training team, who manually updates participant data in the database
End-of-training questionnaire	<ul style="list-style-type: none"> Percentage of graduates satisfied with training Self-assessed changes in knowledge and skills 	Online questionnaire sent to graduates upon completion of the training to measure participants' satisfaction with the training, evaluate initial skills and knowledge gained, and understand participants' short-term plans for looking for work. There are 20 questions, taking 10–12 minutes to complete
Two- and four- month follow-up surveys	<ul style="list-style-type: none"> Percentage of graduates applying skills/knowledge Percentage of graduates having found a job Number of recent job applications Number of recently granted interviews 	Online follow-up surveys sent to graduates two and four months after the training ended, with the objective of following graduates' job-search activities and outcomes. There are 20 questions, taking 10–12 minutes to complete, which address the use of different job-search techniques, time invested, applications sent, interviews granted, assessment of FJIJ contribution, etc.
Focus groups	<ul style="list-style-type: none"> Assessment of FJIJ contribution for several indicators 	Group discussions moderated by EFE Maroc's M&E manager, bringing together 8–10 FJIJ graduates, involving 4–6 key questions used to complement quantitative findings from online surveys. Moderator is accompanied by a note-taker who records all conversations. Held 2–3 times per year, each lasting around 60–120 minutes

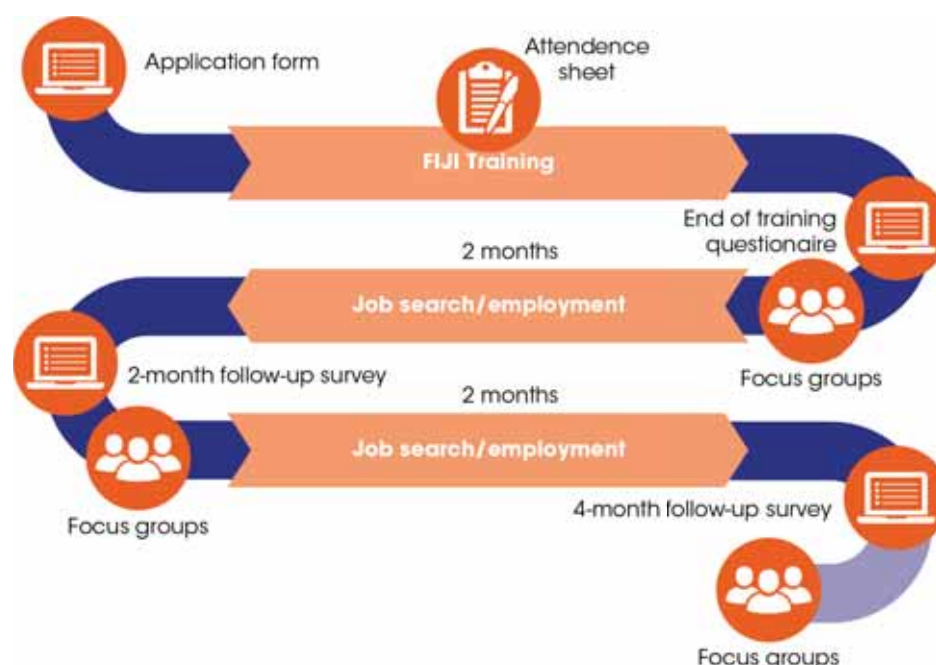
Data from the online surveys can be automatically exported to the spreadsheet, while data from the attendance sheets are transcribed from paper. The M&E manager is responsible for controlling the quality of the data; this includes the arduous task of matching records from the various surveys to the same individual.

In the future, EFE Maroc is seeking to employ more powerful database software in order to better deal with the growing number of data entries and to gain flexibility in the analysis performed.

M&E results

The data collected provided tentative evidence that EFE is achieving its objectives for FJIJ. However, these results should be interpreted with caution because of low response rates (uncertainty as to whether the results are representative of all graduates) and the lack of a comparison group (causality between the project and the observed outcomes can only be assumed).

FIGURE 2. Overview of data collection workflow of the M&E system for FJIJ



- **Self-assessed knowledge and skills:** EFE asked graduates to assess their level of confidence before and after the FJIJ training regarding different aspects of job-search techniques and skills. Results highlighted in figure 3 show that in the seven topics covered during the training, the graduates' level of confidence upon the completion of the training was higher after the training than before the training.
- **Job-search behavior:** EFE Maroc was also interested in knowing whether participants' job-search techniques changed after training. Findings suggest that participants employed all techniques more frequently after training, though using personal contacts for job search (networking) showed only slight improvement, indicating that EFE Maroc might want to reinforce this portion of the training.
- **Securing a job:** The early findings indicate that about half of the job-seeking graduates found a job within four months. Interestingly, graduates who found a job did not necessarily seem to attribute this success to training. Subsequent focus group discussions could help better understand these results.

FIGURE 3. Before and after comparison of confidence in job-search-related activities, 2012 FJIJ graduates (N = 285)



Lessons learned

Survey random sample of your beneficiaries and ensure high response rates

EFE regularly sends online surveys to hundreds of young beneficiaries per month. All youth who participate in EFE trainings are sent both an end-of-training survey and two follow-up surveys. Response rates to the surveys have often been low; in a recent EFE survey, only 24 percent of surveyed youth provided a response. In the future, EFE will not need to send their surveys to all training participants and instead can send the survey to a smaller sample of respondents, while implementing new strategies to get high response rates from those surveyed.

- Calculating the size of a representative sample can be done by using sample-size calculators available online.¹ The larger the size of the population of interest (i.e., training participants), the lower the percentage of individuals needed for the sample.
- The youth surveyed (i.e., the sample) should be drawn randomly from all trainees. This will ensure that the respondents share, on average, the same characteristics as those not surveyed.
- As a rule-of-thumb, the non-response rate should be no more than 20 percent. Strategies for increasing response rates include offering an incentive, sending several reminders, personalizing communication, and choosing alternative survey tools (e.g., phone calls instead of web-based surveys). To prevent “survey fatigue” among beneficiaries, EFE decided to suspend its two-month follow-up survey.

Choose appropriate database management and analysis software

EFE Maroc relied on Microsoft Excel to compile and analyze data about FJIJ. While Excel provides a simple solution, dealing with growing amounts of data case raise challenges. It also requires data collected via an external online survey platform to be imported independently. Practitioners like EFE should seek a data management and analysis software that allows them to better process the large amounts and different types of data that are being collected. In countries with strong

¹ See for example <http://www.surveysystem.com/sscalc.htm>

“Early findings indicate that about half of the job-seeking graduates found a job within four months

internet connectivity, a cloud-based system might be useful, meaning the data would be available online and could easily be shared with staff and stakeholders.²

Ensure the sustainability of the M&E system

EFE Maroc faced high staff turnover in the M&E function. This meant significant delays in implementing the new M&E system. Practitioners anticipating a similar challenge should ensure that the M&E responsibilities should be shared among all several units in the organization.

Next steps

Piloting focus groups

Organizing focus groups amongst FJIJ graduates has been difficult. Focus groups tend to be labor- and time-intensive for staff, as graduates are often unavailable due other commitments like employment. Graduates are also sometimes skeptical as to why they are being contacted, even though careful precautions are taken regarding the wording of the invitation message (email and telephone). As a result of these challenges, the EFE's first pilot focus group had to be cancelled because no participants showed up, and a second one was conducted with only four of eight confirmed participants in attendance. In the future, EFE plans to overcome this issue by offering a larger incentive to youth.

Revisiting M&E system to capture other targeted outcomes

A second major objective (intended outcome) of the FJIJ program is to build the capacity of partner institutions to better serve youth, using examples from successful EFE programs. EFE is currently developing the results chain and indicators tied to this outcome, which will be followed by the design of a data collection plan.

Further enhancing M&E capacity

The Swiss Academy for Development (SAD), a consultancy firm specialized in monitoring and evaluation, has provided external expertise to EFE in this reform process. Upon validation and full roll-out of the system in Morocco, EFE is looking forward to replicating the model among EFE affiliates in Tunisia, Egypt and Yemen.

Assessing the feasibility of an impact evaluation

Finally, when a functioning and strong M&E system is in place, and if judged relevant according to future plans of FJIJ, an impact evaluation (IE) could be performed on the training program. A study to assess the feasibility of conducting an IE of FJIJ was conducted by SAD. It recommended postponing the evaluation until targeting and admission procedures were revised. SAD also considered it risky from a cost-benefit perspective to evaluate an intervention of such a short (three-day) training activity.

2 Examples of cloud-based data management systems include "Salesforce" and "goProve".



A new day with rich hopes for better future

CASE STUDY 3

HELP LEADS TO HOPE, SOMALIA



How a basic monitoring system can foster learning, accountability, and improve service delivery in a fragile environment

KEY MESSAGES

- The knowledge, skills, and lessons learned from basic monitoring activities provide a solid foundation on which more complex evaluations can be built.
- It is important to keep the M&E system simple and data collection requirements in line with the organizational capacity.
- Pilot testing of data collection instruments is critical to understand local sensitivities in providing personal information.
- For small-scale operations, Microsoft Excel can be used to record monitoring and evaluation data. However, it is important to consider the eventual volume of data, how it will be managed, and by whom.

Background

Help Leads to Hope (HLH) is a non-governmental organization based in Bossaso, Somalia, which delivers humanitarian services to marginalized populations in the Puntland region. In 2012, HLH introduced a one-year program to train 380 young people aged 18–35 in fishing skills called the Fishing Sector Livelihood Program for Youth. Beneficiaries of this program learn one of a range of core skills which form the fishing industry value chain, including artisanal fishing, deep-sea fishing, fish processing and conserving procedures, boat and engine maintenance, and the marketing of marine products. A four-month cycle of theoretical training and practical learning provides youth with the knowledge needed to set up new small businesses. Beneficiaries are paid a small stipend into a bank account set up in their name, which helps them understand banking processes and how to save. Following the completion of training, beneficiaries are offered machinery and equipment needed to set up their own small business.

Taqeem support

When HLH joined the Taqeeem Initiative, the organization lacked a strategy and a plan for monitoring and evaluation. Data were being collected on a few key output indicators, but little outcome information was available. The Fishing Sector Livelihood Program for Youth presented a good opportunity to pilot a comprehensive M&E system given it was a new project and HLH could start from scratch in its development.

Before HLH could start planning the basic framework of a new M&E system, it needed to understand its readiness and ability to conduct M&E. There was a clear mandate from its donors at the Education Development Center and the United States African Development Fund who demanded more attention to results measurement. However, the HLH team had little knowledge and experience with M&E before joining Taqeeem. M&E processes were largely set up to respond to requests for donor reporting. Each HLH activity collected data using different tools, and the organization had no central repository for all the data collected. Moreover, HLH lacked the resources to set up a rigorous M&E system. Based on early assessments, HLH and Taqeeem decided to start by building the M&E skills and knowledge among HLH staff, while injecting technical and financial support into the M&E system to make up for a lack of resources.

M&E system

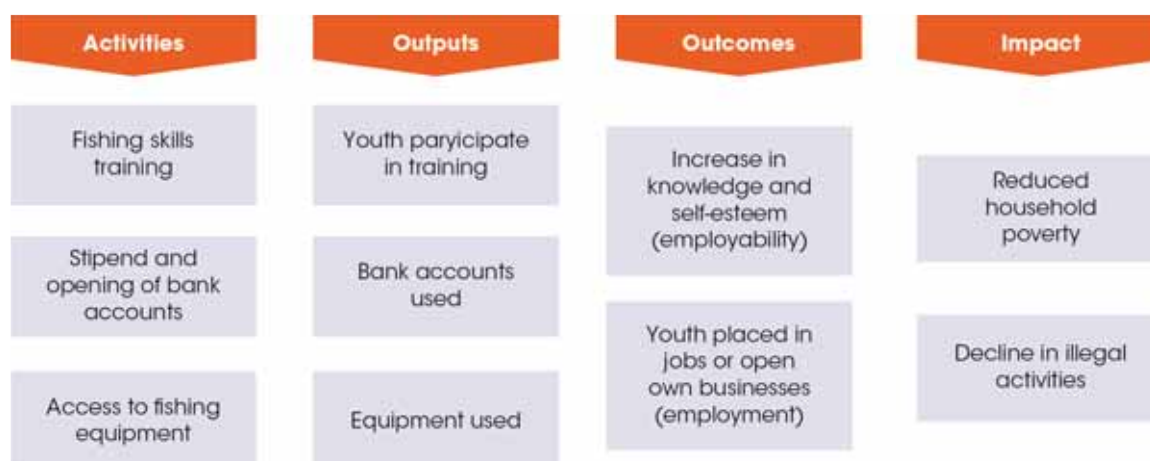
Results chain

The fishing skills program was designed following the old proverb, "Teach a man to fish and you feed him for a lifetime." The aim of the fishing skills program is to teach beneficiaries a new set of skills and attitudes for jobs and business which will provide them with livelihood options, as well as to empower them in terms of self-esteem (figure 1). By providing youth with the opportunity to work, start a business, and earn money, HLH expects to help them move out of poverty and provide basic needs for their families.

Data collection instruments

Working with Taqeeem, the HLH team developed a number of instruments for capturing outputs and outcomes related to the fishing skills program. Given the lack of human resources HLH could

FIGURE 1. Overview of the results chain for HLH's fishing skills program



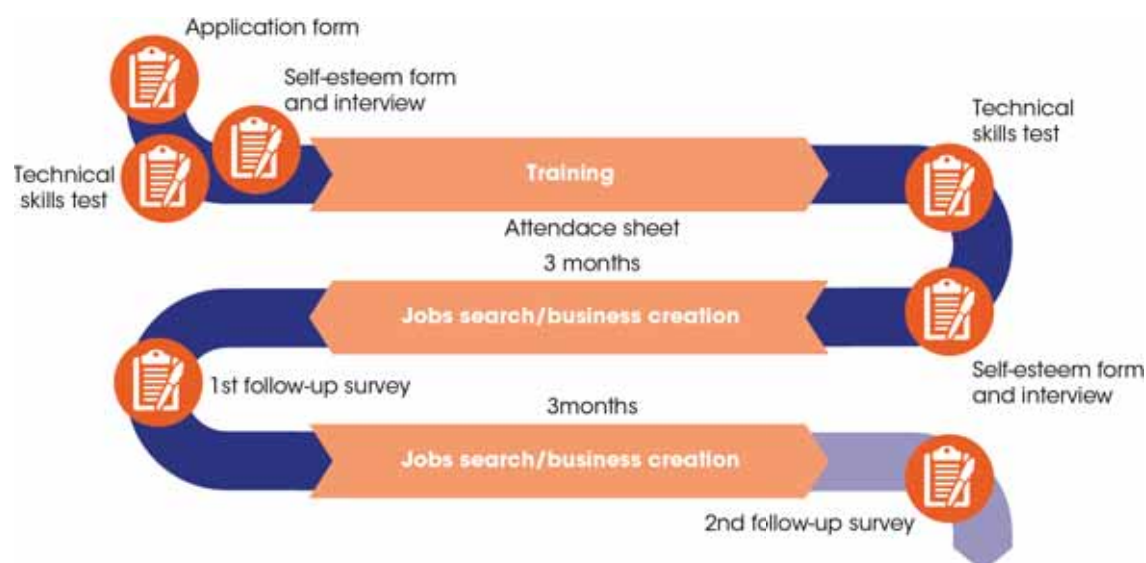
dedicate to M&E, the quantity and the frequency of data collection was kept to a minimum. The data collection tools were designed to be intuitive and fit to the design of the program. Table 1 summarizes the principal data collection tools and indicators they measure.

Figure 2 shows HLH's data collection cycle, starting with the application form (before the training) and ending with the last follow-up survey (six months after the training ends). For a given cycle (cohort) in the fishing skills program, which lasts three months, a full M&E cycle, from beginning to end, therefore lasts ten months. It is worth noting that all data collection instruments were translated from English to Somali to ensure appropriate understanding by the beneficiaries.

TABLE 1. Description of the data collection tools and main indicators of the M&E system for fishing skills

Instrument	Main indicators measured by the instrument	Description
Application form	<ul style="list-style-type: none"> Number of applicants to program Socio-demographic information such as household income, assets, education level and work experience 	Paper application form with 10 questions to capture the socio-economic situation of the applicant. The questions are used to provide an indication of the likely poverty level. Once completed, the M&E Officer is responsible for manually entering data into an Excel data sheet
Self-esteem form and interview	<ul style="list-style-type: none"> Attitude towards employment Motivation to get a job Level of self-esteem 	Paper form completed by the M&E Officer during face-to-face interviews. The form contains 10 statements scored on a Likert scale. ¹ The form is administered before the training starts (as part of the screening process) as well as after training completion. All the data are entered manually into an Excel data sheet by the M&E officer
Pre- and post-training technical skills tests	<ul style="list-style-type: none"> Level of technical knowledge on fishing skills and practices 	A set of five paper-based, multiple-choice questions on technical aspects of fishing covered by the curriculum. All participants take the skills test at the start of the program and again at the end. Results are used to confirm a pass mark to the training
Endline, 3-month, and 6-month follow-up surveys	<ul style="list-style-type: none"> Share of graduates securing a job after 3–6 months Share of graduates starting their own enterprises after 3–6 months Poverty likelihood at endline and after 3–6 months 	Paper forms administered by the M&E Officer at graduation as well as 3 and 6 months thereafter. The same questions are used as in the application form in order to be able to measure change in poverty. Questions on employment status and self-esteem are also asked. The M&E Officer is responsible for entering all of the data on the forms into the Excel data sheet

¹A Likert scale is a widely used approach to scaling responses in survey research.

FIGURE 2. Overview of the data collection workflow of the M&E system for fishing skills

Data management

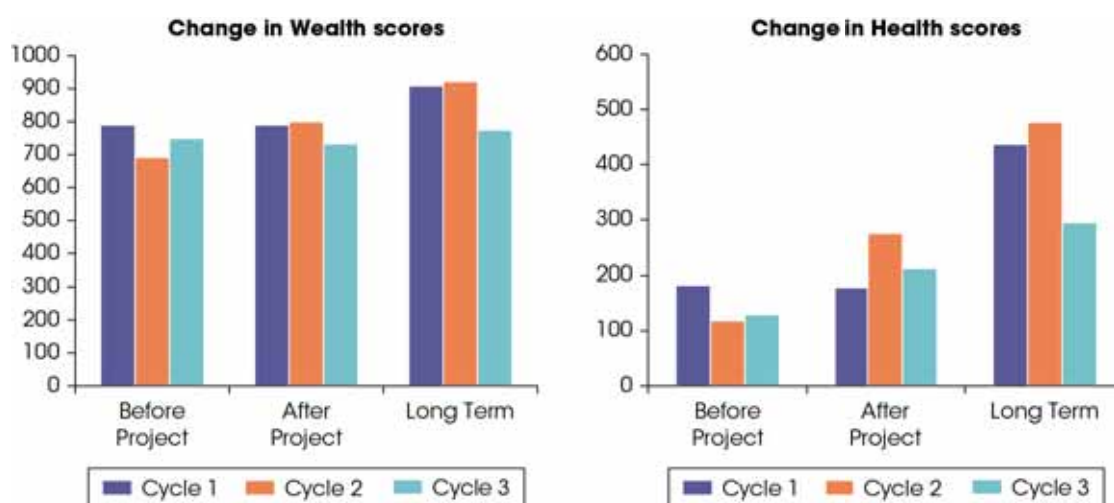
HLH's dependence on paper-based forms and manual data entry in when conducting their M&E activities might be surprising. Indeed, all data are entered manually into a master Excel spreadsheet where analysis is performed. While a web-based or database solution would have been preferable, it was not deemed feasible due to the lack of IT equipment and reliable internet connection.

The spreadsheet has become the backbone of HLH's M&E system. The sheet was set up to automate recording and analysis processes allowing many of the HLH staff to make use of it after receiving training. A series of videos, training manuals, and instructions was created to provide an always-on source of guidance on the use of the tool. As such, the solution provided to HLH has been appropriate to the conditions in which they are operating and has enabled them to establish a useful tool for monitoring program outcomes.

M&E results

The M&E system designed provides some interesting initial results. Three cycles of the program were completed in 2012–13 with a total of 379 beneficiaries participating in the training. Data on these beneficiaries show the following:

- **Gender balance:** 88 percent of participants were male. The imbalance was driven by the focus on the fishing sector which is dominated by men, as well as cultural constraints which only allow women to participate in specific types of training. This situation is unlikely to change in future cycles.
- **Dropout:** The dropout rate was just over 2 percent over the course of the three training cycles.
- **Knowledge:** 100 percent of students passed the post-training skills test.
- **Employment:** 100 percent of graduating students secured employment after the program, either by finding a job or setting-up their own business. This was a huge success, given Somalia's climate not being conducive to learning outside, and the limited number of employers in the area.
- **Poverty:** Data indicate a 26 percent increase in the number of beneficiary households with access to electricity, private latrines, and access to filtered water.

FIGURE 3. Changes in wealth and health status outcomes among beneficiaries

Based on the pre- and post-training survey, positive changes in the beneficiary's wealth and health status could be observed (figure 3). Wealth scores improved by 10–30 percent, while health scores saw improvements of over 100 percent on average. However, since only monitoring data are available without a reliable counterfactual,¹ one cannot be sure that these positive changes are (fully) attributable to the intervention.

Lessons learned

Keep it simple

It is important to align the complexity of an M&E system simple with organizational capacity, considering that collecting too much data will be costly and require a lot of time to analyze. A limited number of data collection tools and indicators proved sufficient for a small organization

like HLH. The original M&E system design at HLH included elements which were not practical when it came to implementation; ideas like offering beneficiaries the opportunity to provide feedback via SMS, or collecting additional data from other household members were never implemented for this reason. The core theme of the M&E design was to provide a credible indication that the program impacted livelihoods and helped beneficiaries move out of poverty. This has been achieved with a reduced set of instruments and basic data analysis.

“Help Leads to Hope (HLH) is a non-governmental organization based in Bossaso, Somalia, which delivers humanitarian services to marginalized populations in the Puntland region

¹ A counterfactual is a theoretical construct that refers to what would have happened in the absence of the intervention; it is usually captured using a comparable control group not benefiting from the program.

Build capacity for data management

The person responsible for managing data should receive appropriate training in data management and analysis. Otherwise, data collection and management will quickly become a burden for the organization. Our experience of HLH shows that analyzing increasing amounts of data in Excel sheets is not easy. The creation of pivot tables and formulas was required to extract the required information and avoid calculating the scores manually. As the amounts of data grew, the formulas and pivot table definitions often required adjustment. These operations required an understanding of the more complex elements of Excel.

Test instruments to account for local sensitivities

When collecting data, it is important to consider local sensitivities, customs and culture in order to ensure that the beneficiary groups feel comfortable giving requested information. As soon as they feel uncomfortable with the questions, they may demur or simply provide inaccurate answers. Pilot testing of data collection tools is crucial to understand what questions will cause a problem, and allows evaluators to adjust the instruments and adapt the selection of indicators before their launch. In the case of HLH, information about employment and self-employment was collected immediately following the graduation, but data collection during the follow-up surveys was difficult because most of the beneficiaries refused to openly answer some of the questions. In particular, beneficiaries did not feel comfortable providing information related to personal income or family affairs. Similarly, the self-esteem form data were not collected after the baseline, and therefore could not be used to analyze potential changes in this regard.

Next steps

Refining the M&E system

HLH's M&E instruments need to be adjusted to account for the cultural and regional sensitivities in answering certain questions and to better fit its learning priorities. In particular, some of the sensitive questions have to be removed from the survey questionnaires and not pursued during focus groups. Priority must be given to the questions that are providing evidence on key outcomes.

Adapting the M&E system to new implementation sites

HLH's original fishing skills program finished after one year of implementation. In its place, a new program focused teaching a broader set of technical and vocational skills to rural youth is to be delivered using the same approach and M&E system, with slight changes to adapt it to the new context. For instance, new questions were added regarding the families' livestock and dairy production. HLH would also like to move to a more robust, cloud-based M&E system if internet connectivity keeps improving.

Leveraging the results

A key objective going forward is to use collected data to assess the performance of each area of the program and improve the quality of delivery. Feedback from beneficiaries will also provide insight into new ways in which the program can deliver additional value. Further analysis of the data will also be used to produce high-level information for donor reports, outreach campaigns, and the organization's website.

Becoming a center of excellence

HLH wishes to become a center of excellence for M&E in the Bassoso region and possibly all of Somalia. Leveraging the capacity building received and experience acquired in building and improving its own system, HLH sees itself becoming an M&E resource base for other organizations. By providing training and a suite of packaged tools, a locally based M&E support facility could have a marked impact on other NGOs operating in the country, enhancing learning and accountability.



CASE STUDY 4

INTERNATIONAL YOUTH FOUNDATION, EGYPT

Improving data quality through optimization of data flow and data collection instruments

KEY MESSAGES

- When designing a technology solution for M&E, it is important to consider end-to-end workflow to ensure that data are collected, managed, and analyzed consistently. Every manual intervention in the workflow increases the risk for data loss, errors, and delays.
- Maintaining contact with beneficiaries post-program is not easy. Youth-friendly approaches, practical communication methods, and incentives for participation should be considered carefully from the onset.
- Training and management of staff performing M&E data collection and analyses need to be designed properly. Third-party M&E responsibilities need to be carefully monitored to ensure data quality.

Background

The International Youth Foundation (IYF) and Nahdet El Mahrousa (NM), an Egyptian NGO, in partnership with the MasterCard Foundation, are implementing Egypt@Work, a four-year, US\$5 million initiative that addresses the need for increased youth employment and entrepreneurship opportunities in Egypt. The program aims to give 10,000 disadvantaged youth increased employment prospects in both wage and self-employment. The program is delivered to young beneficiaries through a range of local implementation partners (LIPs). These partners are trained to deliver the individual projects and to grow their capacity for wider delivery.

Taqeem support

Prior to receiving assistance from Taqeeem, Egypt@Work already had a comprehensive M&E plan with a well-defined results framework, performance measurement plan, 11 different data collection tools, and a risk-mitigation plan. Taqeeem advised IYF mainly on streamlining data collection processes as well as:

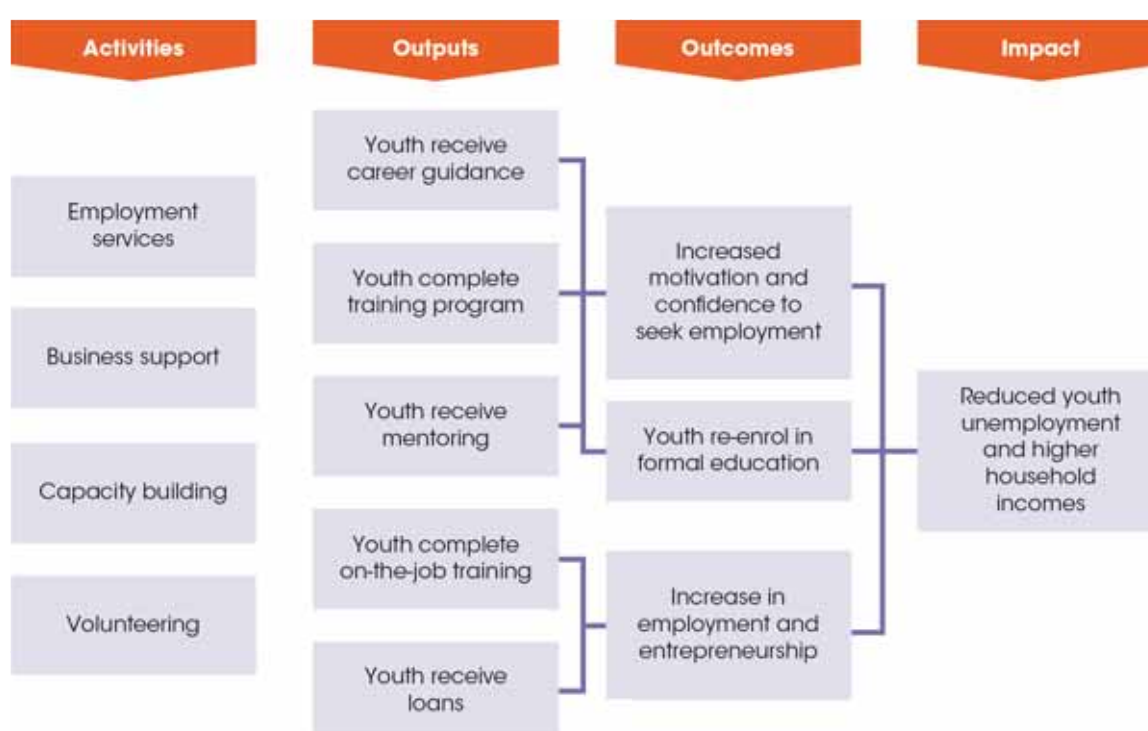
- **Data management:** IYF's challenge was determining how to manage the large amounts of data that were being collected.
- **Capacity development:** Improving the M&E skills and knowledge of IYF staff and partners who were responsible for primary data collection.
- **Survey design:** Improving delivery channels for surveys to increase response rates and accuracy.

M&E system

Results chain

The long-term objective of the Egypt@Work program was to provide young people with market-relevant technical and life-skills training and pre- and post-training support services to increase their employment and entrepreneurship prospects (figure 1).

FIGURE 1. Overview of the results chain for Egypt@Work project



Data collection instruments overview

The main data collection tools of the M&E system are described in table 1, and figure 2 illustrates the data collection workflow.

Data workflow

Before Taqueem's assistance, IYF depended on local program officers to collect data on beneficiaries on paper and to send this information to the local office, where it was entered manually into a database. Results were then sent to IYF's global headquarters, where Excel-based results were migrated into a larger database. Each of these steps provided a source of delay, and data quality was threatened by potential data entry errors.

In order to support the data capture and management of the information from all of the program's nine local partners across Egypt, it was decided to implement a cloud-based platform using the Salesforce.com online database.

Now, data from across IYF's nine implementing partners is consolidated into one online database that can be accessed from any internet connection. The database was designed to hold data

TABLE 1. Description of the data collection tools and main indicators

Instrument	Main indicators measured by the instrument	Description
Baseline survey	Percentage employed in quality jobs or self-employed Percentage who have started loan repayment Percentage in education or training Percentage reporting self-confidence	10-minute, paper-based survey conducted at registration at registration before training begins
Exit survey	Number who complete program Number who receive job or internship placements Number who receive small business support Average satisfaction rates Number of job offers	10-minute, paper-based survey conducted at immediately following graduation
Post-participation discussions and polling¹	Percentage who retain jobs and business Percentage of loan repayment Percentage to re-enrol in education	(1) Focus-group discussions with 10–15 youth (2) A polling method in a group setting Conducted 6 months after graduation
Employer questionnaire (6 months later)	Employer satisfaction with program graduates and participant skills	5-minute, web-based survey conducted 6 months after graduation
Post-program focus groups (12 months later)*	Percentage of new enterprises still operating after 12 months	Survey of youth who had started business at time of exit survey conducted 12 months after graduation

¹Post-participation and post-program surveys utilize sampling meaning that only selected beneficiaries will be surveyed. It is expected that approximately 1,800 out of the total of 10,000 youth will be surveyed

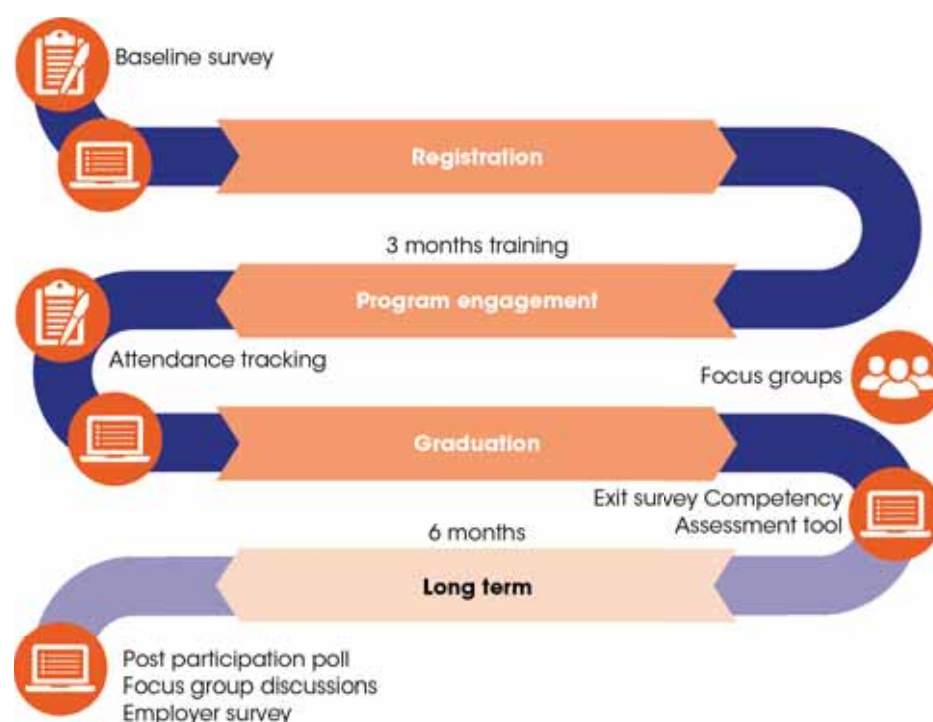
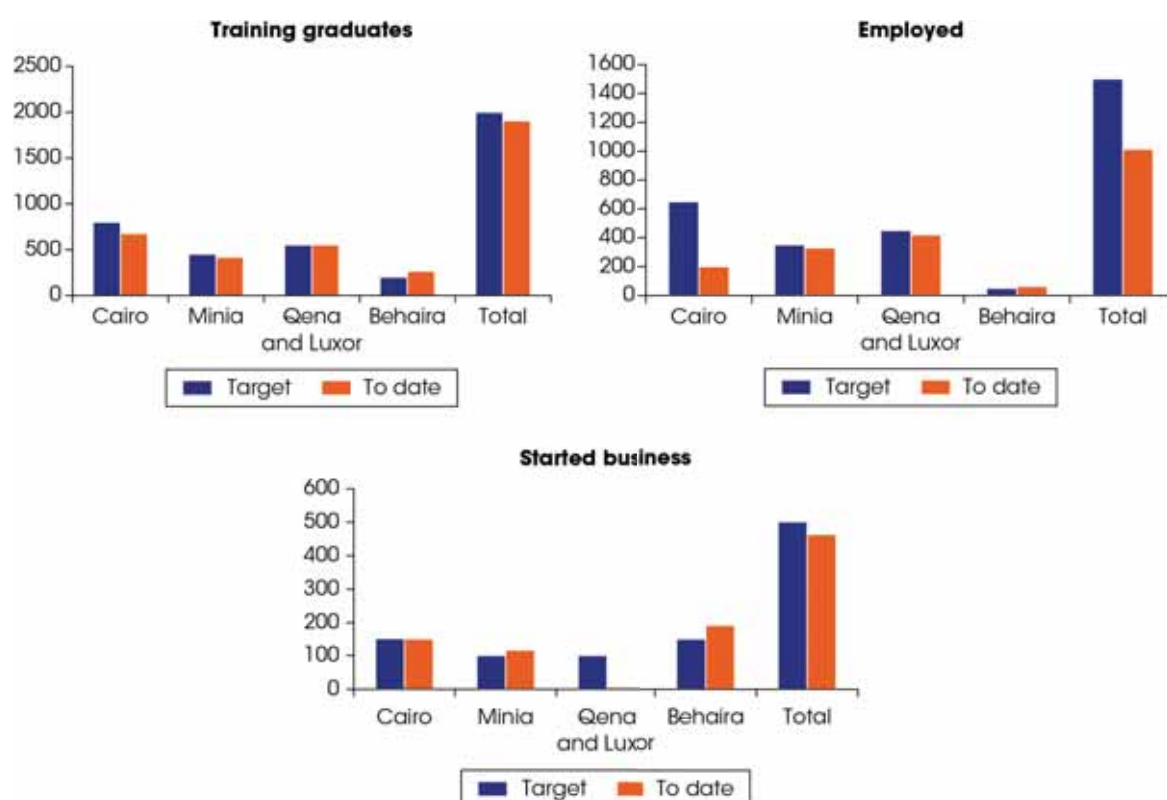
FIGURE 2. Overview of the data collection and intervention workflow

FIGURE 3. Results achieved between January 2012 and September 2013 in four governorates. Targets are based on projected achievements after two years of implementation



from 20,000 baseline and exit surveys (for 10,000 program beneficiaries) as well as other data for the 10,000 youth, including attendance logs and competency results, post-participation surveys and post-program surveys. The system was designed to sort data by individual, partner, sex, and location of training.

M&E results

The results in figure 3, exported from the Salesforce.com cloud, show how IYF uses the results of its M&E system to monitor outcome indicators and track program beneficiaries. In this example, actual figures are compared against target figures to show progression towards committed results. Data from IYF's three main outcome indicators are reported: training graduates, employed youth (wage employment) and business start-up. Data were obtained through post-participation polling 6 months after training graduation and are presented by separating the data from the four governorates where IYF operates.

Lessons learned

Developing and deploying the cloud-based database

While the adaptation of the cloud-based database helped to streamline the data collection and analysis processes, IYF continued to struggle with data quality issues due to a lack of ICT skills in their implementing partner offices. Primary data are still collected using paper forms and data get lost in the transfer from paper to online database. Moreover, the adaptation of the technology itself imposed costs on IYF in terms of staff training and high annual subscription fees.

“Egypt@Work addresses the need for increased youth employment and entrepreneurship opportunities in Egypt. The program aims to give 10,000 disadvantaged youth increased employment prospects in both wage and self-employment

Despite challenges with the data collection, the online cloud-based database is a powerful tool that offers unique advantages to Egypt@Work. Unlike the experience with many other databases, IYF has not experienced any glitches with the system that would require data to be re-entered or checked for accuracy. Likewise, Salesforce.com can hold an extensive amount of data (over 20,000 surveys with 30-50 questions per survey). The database is also a real asset to the Egypt@Work program as it is relational, meaning that IYF can link youth to the implementing partners that provided them service, or link the baseline and exit surveys to the same individual to quickly see changes in individuals over time.

Database adaptation and maintenance can be a burden

The M&E database using Salesforce.com requires complicated coding to adjust survey forms as well as setting up “dashboard displays” (visual graphs showing progress towards goals for each indicator). IYF ended up investing extra funds to hire an IT expert to make the required adjustments.

Challenge of tracking beneficiaries post-training

Six months after youth graduate from the program, IYF calls the youth to determine their job or business status. It was discovered that many of the phone numbers no longer worked, and many of the youth did not appreciate receiving calls from IYF’s central managing partner, with whom they were not familiar. Thus, IYF has started setting up focus-group discussions through the local implementing partners to administer the follow-up surveys with a smaller sample of graduates. Arranging focus-group discussions has also proved difficult. Many youth did not show up for scheduled meetings due to the fact that they were working or in school.

Pay attention to survey design to increase the reliability of responses

It is important to ensure that the questions posed in any survey allow the respondent to give the most appropriate answer. This is a common problem with online surveys, or sheets offering Likert scales for all questions. The temptation to simply tick all questions in the “agree” column is great and shortens the time spent completing the survey. However, the data collected will be of little value and will not reflect the true story. To avoid a fully positive or fully negative response, questions can be phrased to require a negative response in a few cases to indicate a good outcome.

Next steps

Egypt@Work will continue administering the baseline (registration) survey and exit surveys to all youth through M&E Managers based in implementing partners. The surveys will continue to be administered in an interview style to meet the needs of illiterate youth in the program. Data from these forms will be entered into the Salesforce database as youth graduate from the program. This data entry and analysis will be completed by local staff in Egypt who can quickly follow-up with local partners as needed.

A large evaluation of IYF's life skills training program will be conducted in 2015.

As Egypt@Work continues piloting M&E tools and systems, the program will analyze the information collected from the program, partners, youth, and employers about what works in skills training in Egypt. Emphasis will be placed on sharing this knowledge with local partners to help them monitor, improve, and ultimately sustain their work beyond the Egypt@Work program.



CASE STUDY 5

LOTHAN YOUTH ACHIEVEMENT CENTER, KUWAIT



An NGO in Kuwait designs a simple and scalable M&E system through participative learning and harnessing the powers of technology

KEY MESSAGES

- Introducing new technology in an organization that has not yet “digitized” needs to be complemented by thorough change management plan.
- The introduction of monitoring and evaluation into an organization must involve all staff to ensure that the body of knowledge created is understood by everyone.
- Data collection activities should be included into the delivery processes of every project so that it becomes part of the way of working.
- Ensure Board members, trustees, and senior managers understand the value of monitoring and evaluation.

Background

LOYAC is a leadership organization that runs programs for young Kuwaitis aged 6–28 years. One of LOYAC’s strengths is its relationships with over 100 private sector companies that support them financially and provide LOYAC’s graduates with internships.

The flagship program of LOYAC is the Summer Program that runs for six weeks every summer. This program serves 500 youth per year and is accredited by two national universities. The three main components of the program are:



1. **Paid internships** – to provide professional skills to youth through on-the-job training in the tourism, customer service, restaurant, and retail industries (30 hours per week for five weeks).
2. **Community service** – to develop young people's sense of humanity through community volunteering (16 hours total over the six weeks).
3. **Recreational and cultural learning** – students work in teams to create a group project to learn about their culture, explore their creativity, and display their talents.

Taqeem support

LOYAC joined Taqeem to understand how M&E could assist them in proving the impact of their programs to donors and supporters and enhance efficiency in the way they delivered their programs. Prior to engaging with Taqeem, there was no M&E system in place at LOYAC. Very few data were being collected at the output level (e.g. time spent in training and volunteering) and data at the outcome level (e.g. graduation and job placement) were not collected systematically. The understanding of the importance and benefits of solid M&E was low.

Taqeem assisted LOYAC to build its system from the ground up. Much of this work was on capacity development, training LOYAC staff on key M&E concepts. Three LOYAC staff attended trainings, experts made multiple site visits, and consultations with senior management took place.

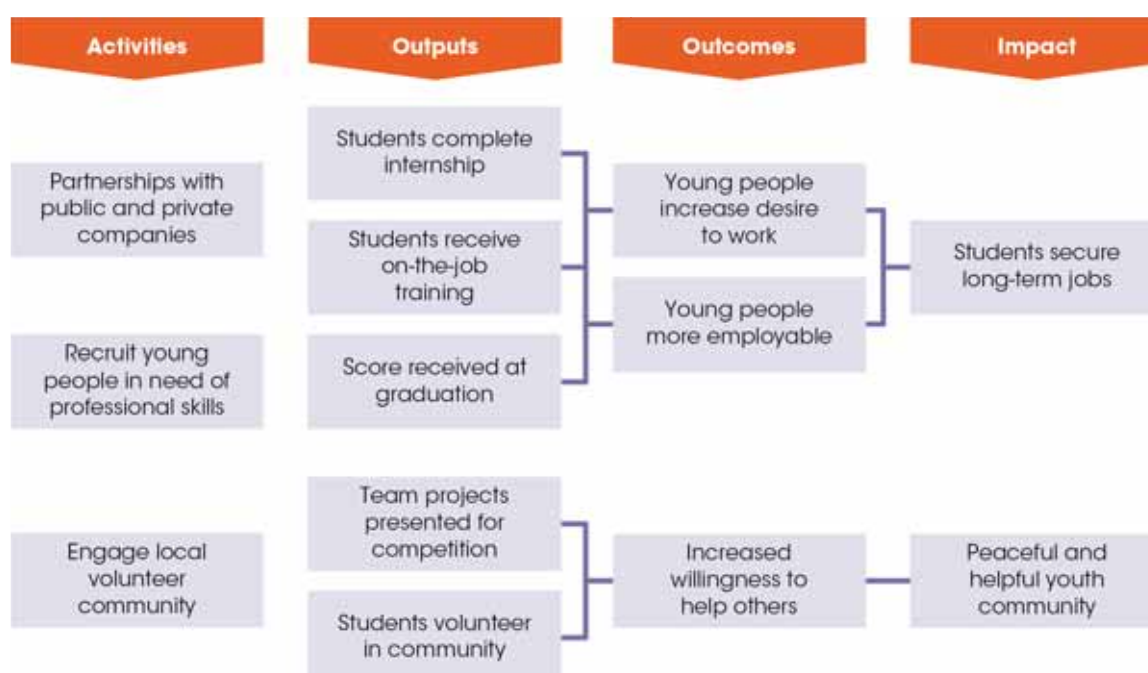
A second area of support was data management. LOYAC was using a basic Management Information System (MIS) to manage student applications forms (name, contact details and basic demographics), and to record who was assigned to which company for the internship. A newly designed cloud-based M&E system was introduced called GoProve. Survey data are now collected at several points during the program from all program participants (see table 1).

M&E system

Results chain

The results chain for the Summer Program summarizes LOYAC's theory of change for addressing a lack of employable skills among Kuwaiti youth as well as a lack of understanding of importance of being involved in the local community (see figure 1). In the long run, LOYAC wants the internships to lead to careers in the companies where they are placed and for youth to run successful community projects.

FIGURE 1. Visualization of the results chain from activities to impact at LOYAC



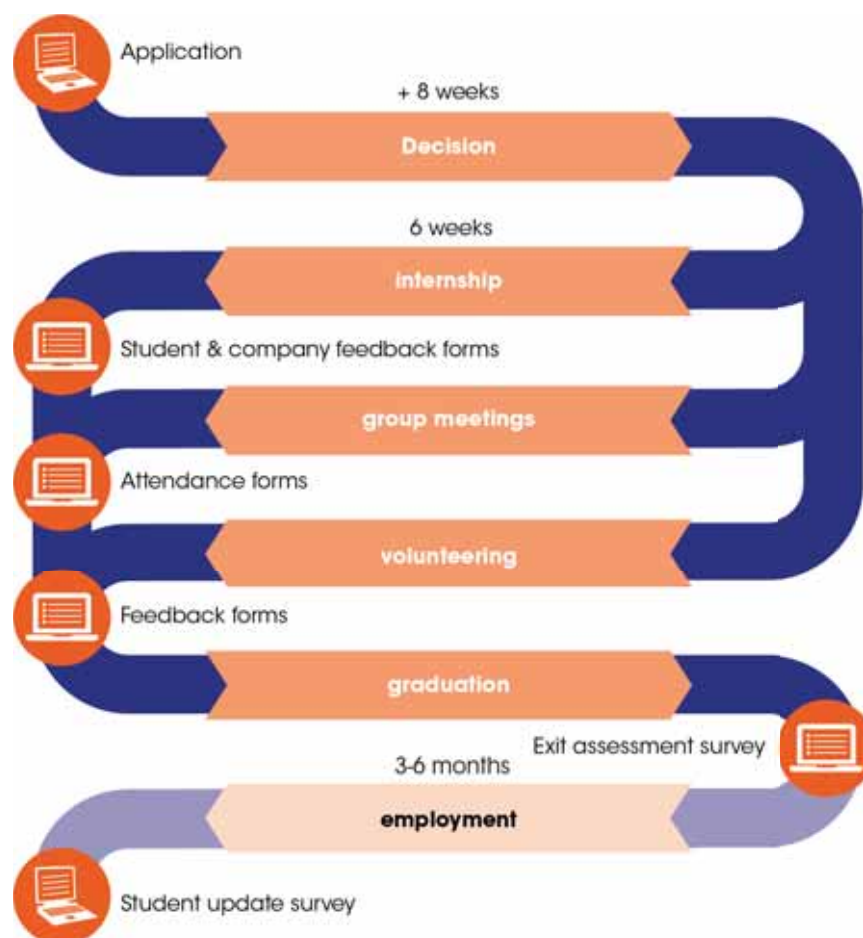
Data collection instruments

The main data collection tools of the M&E system are described in table 1, and figure 2 illustrates the data collection workflow.

Data management

Taqeem supported LOYAC in upgrading their M&E database to a simple, more intuitive system *goProve*, a cloud-based application designed to collect, manage, and analyze M&E data.¹ The application enables data to be fed by a series of survey forms. For instance, *goProve* provides the ability to employ offline data collection tools using tablet devices. Similarly, the students could also send in feedback via email directly to the *goProve* database. Through the new system, data can be collected by LOYAC supervisors at the office, at company premises, and at other locations throughout Kuwait. Moreover, data from the existing MIS system could be uploaded into the new database. In terms of analysis, *goProve* allowed for live analysis of the data, so dashboards were configured to provide LOYAC's management with information on data collection progress and results analysis. Analysis, charts, and reports were used to present to the Board at regular intervals.

¹ For more information on *goProve*, see <http://www.goprove.org>.

FIGURE 2. Overview of the data collection workflow**TABLE 1.** Description of the data collection tools and main indicators

Instrument	Main indicators measured by the instrument	Description
Application form	Demographic information: gender, education, work, and volunteer experience	Registration forms are completed online through LOYAC's website and data are stored in the cloud-based database
Application form Exit assessment survey Update survey	<ul style="list-style-type: none"> Percentage of students to increase professional competency scores Number with intention to seek work in private sector Number who intend to continue volunteering Percentage Increase in social awareness scores 	<p>Self-reported or supervised questionnaires to assess student's professional competency, motivation to work and social awareness using a five-point Likert scale</p> <p>The short questionnaire is completed three times by all students: at the application, immediately following graduation and three to six months following graduation. The survey data from each student are automatically reconciled and linked to the student record by goProve, allowing individual and disaggregated analyses on the changes in response over time</p>

(continued)

TABLE 1. (continued)

Company evaluation form	<ul style="list-style-type: none"> ■ Professional competency score ■ Social awareness score ■ Attendance score ■ Attitude score ■ Appearance score ■ Teamwork score 	Completed by the host companies at the end of placement to provide feedback on student's performance using a five-point Likert scale. Completed online through the goProve website
Feedback form	<ul style="list-style-type: none"> ■ Student's satisfaction with supervisor, job placement and LOYAC support ■ Employer's satisfaction with student and LOYAC support 	A series of eight questions scored on a five-point Likert scale provides a way for host companies and students to give quantitative feedback on their satisfaction with the program. Feedback can be given and submitted at any time through the M&E database

M&E results

M&E data were obtained from the 2013 cohort of summer students where 477 students were accepted into the program.

Baseline characteristics

Overall, the program attracted male and female participants equally. Fifty percent of participants were either in or graduated from high school, while 40 percent were pursuing higher education (see figure 3).

Exit assessment survey

Of the 477 students, 77 percent completed the exit survey providing insight into the success of the program in increasing employment competencies and social awareness. Figure 4 shows data from three of the main survey questions.

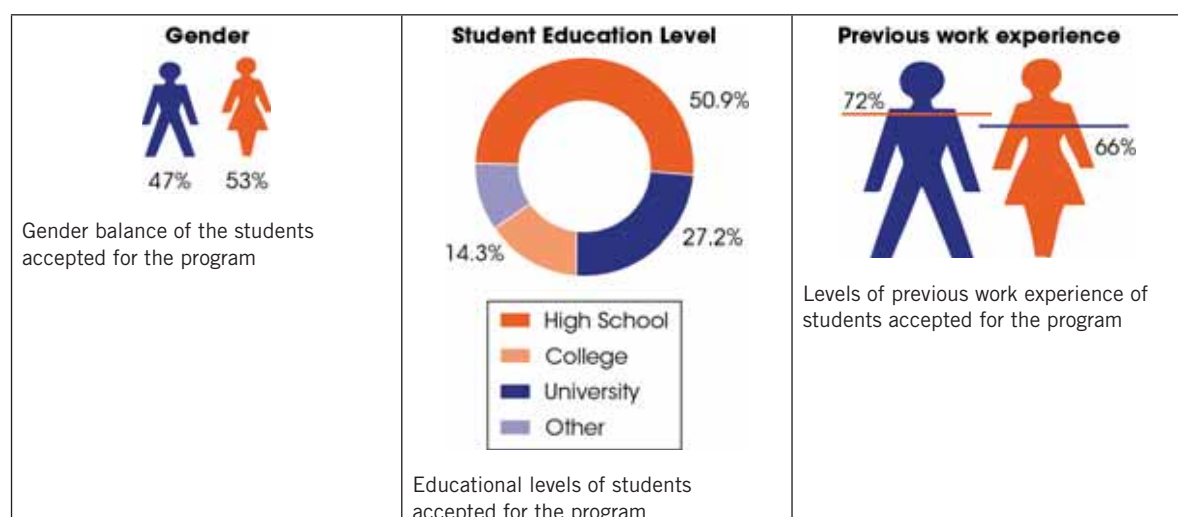



FIGURE 3. Baseline demographic information

FIGURE 4. Exit survey data based on 367 student responses

Do you know the hiring process of at least one private sector company?  44%	Do want to get a job as soon as possible when you graduate even if this is not your dream?  60%	Do you always treat all of your friends with respect?  42%
---	---	---

Lessons learned

Ensure buy-in by senior management

For a change process towards stronger M&E to be successful, the Board and senior management team need to understand the benefits of collecting data and being able to demonstrate their programs' effectiveness and provide an enabling environment for implementing stronger M&E across the organization. More and better data about program effectiveness, in turn, can act as a powerful tool when engaging with prospective delivery partners and funders.

Embed M&E knowledge in more than one person

The training and support was focused on the appointed M&E Officer within LOYAC. Soon after the start of Taqueem's support, the M&E Officer left LOYAC and the knowledge and expertise went with him. The lesson learned here was to ensure that more than one person receives training and support. As people get trained in new techniques, they become attractive to other organizations that will tempt them away, leaving a knowledge gap. In response, all LOYAC staff and volunteers have now been provided with new job descriptions.



Keep it simple

Initially, the M&E system proposed capturing video feedback from a selection of the students at the start of their engagement (interview) and then at the end to gather a series of change stories as told by the students themselves. The aim was to create a series of short before- and-after stories that could be used to promote the program. Due to the severe resource issues, some initial videos were planned, but were not taken any further. The lesson here is to make the first implementation of any M&E system

as simple and lightweight as possible. This will encourage data collection, make analysis easier, and produce simple reports with little impact on delivery processes. Over time, the detail of the data and complexity of analysis can be increased.

“LOYAC’s Summer Program provides professional skills to youth, develops their sense of humanity, and helps them learn about their culture, explore their creativity, and display their talents

Make it relevant

It is important to collect data from the students during the program or very shortly afterwards so that the response rate is high. Attempting to request data a long time after the end of a program will be much more difficult. A strategy that may help to collect follow-up survey data is to educate the students on why the data are being collected and the benefits to the program. This “giving back” will appeal to many of the students, but should not be seen as a guarantee.

NEXT STEPS

Adopt M&E for other programs

The M&E system designed for the Summer Program has been adapted slightly to meet the aims of other LOYAC programs, such as the internship program. The main difference from the Summer Program is that there is no volunteering element; however, the desired outcomes remain essentially the same. By keeping the same set of questions for the majority of the programs, LOYAC should be able to monitor the change in those students that attend multiple programs.

For new iterations of the Summer Program, the M&E system will be simplified so that only the essential data and surveys are collected to monitor some of the important outcomes. Over time, additional indicators and surveys will be added as resources and stakeholder demand changes.

MIS versus M&E

LOYAC has a comprehensive MIS database as well as two permanent members of staff who design, build, and maintain this system. The LOYAC strategy is to define and build new functions into this system to allow their partners (companies) to enter data remotely. The data required for M&E reporting and analysis are exported from the MIS and imported into the M&E database, providing a quick and simple way to share the data without disrupting the proven MIS processes.



Using the results

Prior to participating in the Taaqeeem initiative, LOYAC had no results information available and did not systematically use the information collected from the MIS. Moving forward, it is intended to use the data collected more strategically to assess performance and improve program delivery, and leverage the information in communication and outreach to donors and the public (e.g. reports, website).



CASE STUDY 6

MENNONITE ECONOMIC DEVELOPMENT ASSOCIATES, MOROCCO



How a mid-sized NGO is partnering with external researchers to carry out the first impact evaluation on youth employment in the country

KEY MESSAGES

- The case of MEDA Maroc provides good evidence that close involvement of the implementing organization in the impact evaluation facilitates the implementation of the research.
- Mobile data collection equipment used efficiently can substantially decrease the required time and resources and increase data quality.
- Ensuring high take-up of the program is key to a successful study.
- Impact evaluation studies can be small and localized, although tests to determine the extent of external validity should be run before the study begins.

Background

MEDA Maroc and the “100 Hours to Success” training program

The Mennonite Economic Development Associates (MEDA) Morocco implements the YouthInvest project, a five-year initiative sponsored by the MasterCard Foundation. The objective of the project is to provide economic opportunities for Moroccan youth through skills and employability training, on-the-job experience, access to finance, and providing support to NGOs and financial service providers in the development of financial and non-financial products. The project comprises four different training programs, the largest one being “100 Hours to Success.”

“100 Hours to Success” is a program aimed at youth between 15 and 25 years. It provides them with 100 hours of training in life skills, entrepreneurship, and financial education. The participants of the training are required to open a bank account at Al Barid Bank.¹ After the training, between 5 and 10 percent of youth receive internship placements. The length of the training program is three months. Since 2009, approximately 20,000 youth in Morocco have benefited from the training.

¹ MEDA Maroc negotiated the minimum deposit amount down from 100 Moroccan dirhams (MAD) to 5 MAD.

management role. An expert investigation team was assembled; it was responsible for the scientific validity of the study's design as well as the analysis of the data.

Impact evaluation

Design

The study, being conducted in the oriental region in the cities of Oujda, Jerrada, and Taourirt, is the first experimental evaluation⁶ focused on youth employment to be conducted in Morocco. The evaluation seeks to determine the impact MEDA Maroc's "100 Hours to Success" training program on youth labor market, education, and savings outcomes.

The research team chose a randomized controlled trial (RCT) design for the evaluation. In RCT designs, surveyed individuals are randomly divided into a participant and a comparison group. Differences in outcomes between both groups are calculated with the only assumed difference between the two groups being the intervention itself. The study involves a baseline survey (conducted in October and November 2012) and a follow-up survey conducted 12–14 months after the end of the training.

Baseline data collection

An external data collection firm, made up of 16 enumerators and three field supervisors, was hired to conduct the data collection. The questionnaire contained between 100 and 122 questions, and a full questionnaire took about 25–30 minutes to administer. The enumerators were trained in survey techniques and the survey instrument was tested with a small group of youth and subsequently adapted to ensure an exhaustive understanding by prospective survey respondents.

The surveys were conducted in 13 youth centers where MEDA Maroc's trainings are carried out. Youth were invited to the center which was closest to their place of residence. Enumerators carried out the surveys using tablets equipped with Open Data Kit (ODK) survey software⁷. Questionnaires were uploaded onto the tablets, and responses were then collected electronically. Completed surveys were uploaded onto a central server and responses were aggregated into a format ready for analysis.⁸ This mobile data collection technique saved the team valuable time and resources and greatly reduced the number of errors usually caused by manual data entry. The survey was complemented by four focus groups.

Randomization and take-up

Over the three-week span of the baseline survey, 1,817 youth were surveyed. A random selection initially placed 600 in the participant group (equaling the number of available spaces in the training program during the time under study) and 1,217 in the comparison group. The bigger size of the comparison group was decided in order to both increase the power of the calculations (given the fairly small population) and help remediate some of the enrolment take-up issues that

6 Various experimental evaluations on youth training, but conducted in other countries, can be found in the Youth Employment Inventory (<http://www.youth-employment-inventory.org/>), such as *Program for the Promotion of Children and Youth* (Uganda, 2006), and *Tap & Reposition Youth – Savings & Micro-credit for Adolescent Girls* (Kenya, 2005).

7 ODK is open-source software that helps organizations author, collect, and manage mobile data collection.

8 The complete datasets are finalized on Excel.

MEDA foresaw at the time (furnishing back-up candidates for the participant group). Summary statistics show that randomization was successful in achieving balance across participant and comparison groups.

Nevertheless, despite considerable efforts, the take-up rate of the training proved troublesome: of the 600 youth who were invited to the training, only 340 were in attendance by December 2012. To achieve their objective of 600, MEDA decided to organize a second wave of training. Therefore, a second random selection of 300 youth was taken from the comparison group and brought into the participant group, bringing both into alignment at about 900 individuals each. The second wave of training was completed in August 2013, with a combined total of 505 youth having participated in the training overall.

Baseline findings

A baseline survey taken alone is insufficient to demonstrate any changes on outcomes, not to mention impact, since the treatment has not happened yet. That said, the baseline report provides an extensive descriptive analysis of the population under study. Key findings from the baseline survey are listed below.

Gender balance

The surveyed population of 1817 youth is 53 percent female. As shown in table 1, male program applicants are younger than their female counterparts. The observed age differential by gender is probably related to higher educational attainment rates by young women in Oujda (and Morocco), while young men are more likely to leave school and seek out employment (and training support for their employment search) at an earlier age. The fact that young women are well represented among participants is a positive finding, since traditionally they are much less likely to engage in activities offered through often male-dominated youth centers.

TABLE 1. Age categories by gender (%)

Age	Male	Female
Total	100	100
15–18	37.6	26.7
19–21	35.7	44.1
22–24	21.2	23.5
25–30	5.9	5.7

Educational attainment

The relative young age of the sample is reflected in the latest level of education attained, as shown in table 2, with 91 percent of the sample having not yet obtained a higher degree than the Lycee.⁹ With the majority of participants enrolled in some form of professional education (30 percent) or university (40 percent), it is also clear that the program primarily attracts a better educated cohort than the average Moroccan youth population. The analysis also features the data broken down into the participant and the comparison groups,¹⁰ showing no statistically significant differences between both groups.

⁹ Lycee corresponds to the last degree before entering university.

¹⁰ Randomizations presented here refer to the second wave of randomization, the one where both groups align to approximately 900 individuals each.

TABLE 2. Attained level of formal education, by groups (%)

	Primary	College	Lycee	Professional degree	University and above	Total
Total	18.6	29.5	43.1	4.9	3.7	100
Comparison	20.1	28.4	42.2	4.7	4.7	100
Participants	17.2	30.5	44.1	5.2	3.1	100

Socio-economic status

The analysis of the household income, also broken down into both groups, shows that the participant group is slightly biased towards poorer populations, as shown in table 3. The share of those in the highest two income brackets constitutes 7 percent of the participant group compared with 10 percent in the comparison group. While the difference between these two groups is not statistically significant (as per the baseline report), future assessments will have to be mindful about this potential bias. Overall, 83 percent of the individuals from the sample (26 + 40 + 17 percent) declare coming from households that have an income smaller than 5,000 Moroccan dirhams per month, confirming that the program also reached lower-income households.

TABLE 3. Household income status (in Moroccan dirhams),¹¹ by groups (%)

	0-1 500	1 500-3 500	3 500-5 000	5 000-7 000	7 000-10 000	10 000+	Total
Total	25.8	40.0	16.9	8.8	4.3	4.2	100
Comparison	25.7	38.6	16.6	9.1	4.9	5.1	100
Participants	25.8	41.4	17.2	8.5	3.7	3.3	100

Employment status

Finally, table 4 looks at the current employment status and illustrates that the vast majority of the sampled individuals are still enrolled in some level of education.¹² The education status of the majority should be kept in mind when evaluating employment outcomes: while it was expected on evaluation design that individuals would be enrolling in MEDA Maroc's 100 Hours to Success as a bridge in the transition from school to work, the actual transition might take longer for many than expected in that a year from now many might still be in school and only passively searching for employment.

At baseline, 85 percent of respondents were not active in the labour market (most of them being full-time students). Out of the active individuals (15 percent of the total), 32.3 percent were employed and 19 percent were self-employed. The unemployment rate among this – albeit small – population of active individuals was nearly 49 percent. Although the participant group is slightly more inclined to be active in the labor market, both groups are largely in alignment, with no statistically significant differences.

¹¹ As of July 2013, 1000 MAD roughly converts into 116 US dollars.

¹² Despite the fact that most youth are still in school, 58 percent of sampled individuals claim having already somehow worked in the past.

TABLE 4. Current employment status, by groups (%)

	Employed	Self-employed	Unemployed	Inactive	Total
Total	4.9	2.9	7.5	84.7	100.0
Comparison	4.3	2.5	7.3	85.8	100.0
Participants	5.5	3.3	7.7	83.6	100.0

Lessons learned

Involvement of the implementing organization

The “100 Hours to Success” study is a rare case of an implementing organization being actively involved in the impact evaluation processes, although not ultimately responsible for the study, which falls under the responsibility of the investigation team led by academics and professional impact evaluation experts. In many cases, the implementing organization is left aside from the study. The comparative advantages of having the implementing organization participating actively in the work are many, and include:

1. **Matching project design to data collection cycle:** The timing of data collection needs to coincide with the intervention which is dependent on the work plan of the implementing agency.
2. **Increase take-up:** A high dropout rate can undermine the validity of any study. By carrying an extensive outreach campaign (phone calls and home visits) and making training schedules extremely flexible, MEDA was able to increase the take-up.
3. **Trust with beneficiaries:** An implementing organization has built-up trust in the communities where they are working and can assist with data collection, especially when trying to collect sensitive information.

Capacity-building in impact evaluation

Many organizations lack expertise and experience in impact evaluation. As in the case of MEDA Maroc, the capacity of implementing organizations to engage in impact evaluations needs to be built. In order to better understand the mechanics of the evaluation, the counterpart at the implementing organization should receive basic training on impact evaluation¹³ ideally complemented by continued coaching to help identify impact evaluation experts and assist throughout the preparation process. Even though an impact evaluation might be mostly conducted by a team of experts and an independent data collection firm, a successful study would not be possible without an enabling environment and the understanding and inputs of knowledgeable focal points in the implementing organization.

Technology

The MEDA Maroc example has shown how the use of mobile technology can significantly increase data quality and reduce costs. The high price of buying mobile devices was well worth the investment. That said, practitioners should reserve enough time to pilot the survey and train the enumerators. The ordering, programming, editing, and connecting of the devices is a time-consuming task.

13 For instance YEN's Evaluation Clinics (<http://www.ilo.org/public/english/employment/yen/whatwedo/projects/clinics.htm>) and J-PAL's Evaluation Courses (<http://www.povertyactionlab.org/course>).

“The fact that young women are well represented among participants is a positive finding, since traditionally they are much less likely to engage in activities offered through often male-dominated youth centers

Dealing with take-up and attrition

High dropout rates are quite common in training interventions. This can introduce bias in the results as those youth that choose to complete the training can be quite different from those that drop out. MEDA faced this challenge during a first wave of training when 38 percent of selected youth did not show up. Trying to remediate the problem, they decided to organize a second wave of training and randomly moved 300 youth from the comparison group to the participant group. While this random selection should help reduce bias in

results, certain heterogeneities in the results will not be observed because of decreased statistical power. At the same time, MEDA has been tracking down dropouts to ensure they can be available for the follow-up survey where “intention to treat” analysis can be applied.¹⁴

Next steps

Tracking youth until the follow-up survey

Given the highly mobile nature of youth in Morocco, there are risks of youth moving away from the program area and changing telephone numbers and email addresses. As such, MEDA Maroc will conduct an SMS tracking survey in order to maintain accurate contact details for both the participant and the comparison group. The exercise will consist of verifying their contact information (address, email, phone) and receiving an update regarding youth employment status mid-way through the study.

External validity

As the study is concentrated in the oriental region of Morocco, there is a concern regarding the external validity of the results. In order to observe whether the findings can be applied to the rest of the country, MEDA will compare their demographic and socio-economic baseline results with the 2009 Moroccan Household and Youth Survey¹⁵ conducted by the World Bank.

Follow-up survey

Originally planned for December 2013, it has been postponed to the start of 2015 because of the addition of a second wave of trainings. Since a follow-up survey in stages was considered too costly, the study team decided to conduct the follow-up at the same time for both training cohorts. While this implies that there will be less time for training impact to materialize for those youth that started the training later, it is expected that these differences will not strongly affect the results. Data analysis will be carried out separately for both cohorts, allowing for disaggregated findings.

¹⁴ “Intention to treat” analysis estimates the impact on those intended to be treated rather than those who actually participated.

¹⁵ See <http://microdata.worldbank.org/index.php/catalog/1546>.



CASE STUDY 7

SPARK AND BIRZEIT UNIVERSITY, PALESTINE

spark
ignites ambition

An overview of a SPARK's experience piloting new data instruments, an M&E database, and a partnership with impact evaluation researchers

KEY MESSAGES

- Technological M&E solutions may save time and effort in the long run but require a large upfront investment and sufficient staff time.
- It is advisable to have an M&E culture and existing data collection process already firmly in place before transitioning to a cloud-based solution.
- Regression continuity designs correlate well to business plan competition programs. Nevertheless, impact evaluation is more challenging than it appears, requiring common understanding on program logic, large sample sizes, and high technical expertise.

Introduction

Founded in 1994, SPARK is an independent, international development organization, aiming to improve higher education and foster entrepreneurship so that young people are empowered to lead their post-conflict societies into prosperity.

In Palestine, SPARK actively supports the Birzeit University IT Center of Excellence¹ through their youth entrepreneurship development program called "Strengthening Technology Entrepreneurship for a Sustainable Palestinian Knowledge-Based Economy," targeting youth living in the marginal areas of West Bank and Gaza. The intervention is structured along three phases. In the first phase, participants attend a basic business knowledge training course and are assisted in developing and drafting a business concept. In the second phase, more comprehensive one-to-one coaching is provided to further refine these business concepts and to develop formally sound and viable business plans. In the third phase, participants are given professional coaches and are matched with appropriate funds to start up their businesses. Participants of the first phase are selected through business plan competitions into the second and third phases of the intervention.

1 Hereafter referred to as Birzeit.

Taqeem support

Prior to Taqeem support, SPARK followed a standard monitoring protocol based on indicators and targets agreed following the Donor Committee for Enterprise Development (DCED)² standard. Data are collected mainly through face-to-face interviews, as well as through e-mailed survey tools. Country Managers monitor data from their country program using SPARK standardized tools and formats.

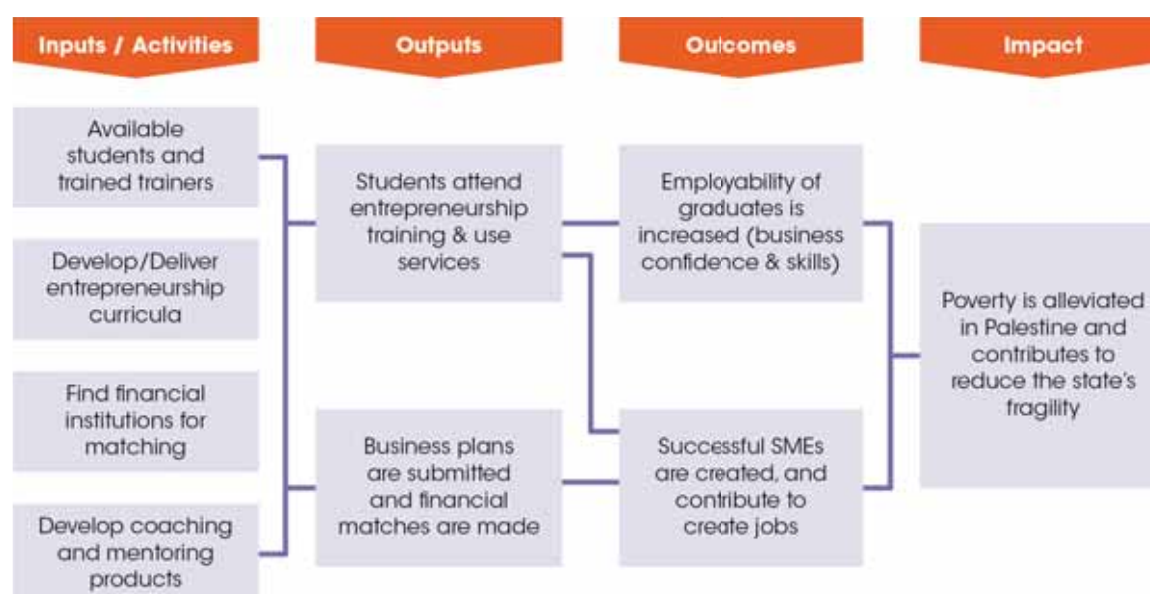
Impact evaluations have been restricted due to resource and time constraints. Due to SPARK's work in post-conflict environments, secondary data are often lacking, and gathering primary data can be logistically cumbersome.

The focus of Taqeem support was threefold. First, to create new and improve existing data collection instruments as well as train staff on the use of the new instruments. These include a longitudinal outcome monitoring survey and a self-recording video tool; second, to replace paper-based monitoring forms with an innovative, digital-based monitoring application. A third focus was to identify experimental impact evaluation approaches to help SPARK rigorously evaluate the impact of their Palestine program.

Results chain

The results chain of the entrepreneurship training program (figure 1) depends on the assumption that jobs are available in Palestine and that financial institutions are willing to support successful business plans. Taking this into account, by delivering entrepreneurship curricula to its young graduates and facilitating the development of solid business plans, Birzeit and SPARK seek to increase employability and self-employment opportunities, which – in the longer term – are expected to contribute to poverty alleviation in Palestine.

FIGURE 1. Overview of results chain, Birzeit and SPARK's entrepreneurship program



Note: SME = small or medium enterprise

2 The DCED standard is a practical eight-point framework for results measurement. It enables projects to monitor their progress towards their objectives and better measure changes, manage implementation, and demonstrate results. (<http://enterprise-development.org/page/introduction-standard>)

Data collection instruments

The main data collection tools of the M&E system are described in table 1, and figure 2 illustrates the data collection workflow.

TABLE 1. Description of the main data collection tools of the M&E system

Instrument	Main indicators measured by the instrument	Description
Application form	Socio-demographic information such as age, gender, income, education	Paper-based application forms collected two months before start of training provide basic demographic information to SPARK which allows them to select eligible youth for the program
Baseline and six-month follow-up outcome questionnaires	<ul style="list-style-type: none"> Number and percentage of graduates with increased business confidence/skills Number of SMEs started Number of jobs created in the SMEs started Number and percentage of graduates finding a job within 12 months Average time needed to find employment 	The quantitative survey was designed to be implemented to the beneficiaries twice; at baseline when youth enter the program and six months after its completion. Comparing both survey results will contribute to measure indicators covering topics such as entrepreneurial behavior change, income increase and job-research improvements. The six -month follow-up survey contains more questions to measure additional indicators such as time needed to find an employment and number of jobs created in SMEs
Performance monitoring tool	<ul style="list-style-type: none"> Number of students completing the entrepreneurship training Number of business plans submitted and financial match-ups made 	This tool aggregates output data on entrepreneurial training delivered, business plan competition winners, matches made businesses started, and jobs created. Means of verification for the data include attendance sheets, business plan competition jury reports, financial agreement documents, business visits, business registrations, and employment contracts or employee details
On-going SME individual tracking for three years	<ul style="list-style-type: none"> Percentage of SMEs surviving after three years 	All SMEs, from their starting date, will be individually tracked for the subsequent three years, to measure the business survival rate

Data management

Taking into account the growing complexity of SPARK's M&E system, it was decided to adopt a cloud-based³ data management platform that would allow Birzeit's/SPARK's business development advisors and staff, regardless of their geographical location, to efficiently enter, manipulate and analyze output and outcome data.

³ A cloud-based system refers to a system made available to its multiple users via the Internet.

FIGURE 2. Overview of the data collection workflow of the M&E system

SPARK decided to adopt the *Salesforce* software⁴ for the business data management component of the M&E system, with the direct consultancy support of *Vera Solutions*.

The system concentrates on business level data including business registration dates, sales, turnover, profitability, and employment. Each entrepreneur has an “account,” managed by Birzeit/SPARK advisors, where data are recorded. The database also serves as a document library, an email client, and an M&E report generator.

Impact evaluation

In designing the impact evaluation, SPARK worked with a group of researchers from the Swiss Academy of Development to propose an appropriate methodology for determining the program’s impact. It was decided early on that an experimental approach (randomized controlled trial) would not be feasible given that program admission was based on self-selection and there was very little excess demand, making it difficult to construct a control group.

The researchers narrowed in on a specific feature of SPARK’s program, the business plan competition, which fits well to a particular quasi experimental methodology: a regression discontinuity design. The in business plan competitions points were awarded based on the technical quality of business plans. In this case, a cutoff score is determined whereby participants who receive scores just above the cutoff (treatment group) are compared with scores just below the cutoff (control group). The difference in outcomes between these two groups tells us if there has been an impact.

There are two main challenges with this type of design. First, since only individuals located in the proximity of the cutoff are included in the analysis, whereas individuals located further away are

4 Known for its customer relationship management system, *Salesforce* software combines database structures and navigation, notably through drag-and-drop reporting and built-in dashboards.

excluded, a large sample size is required. Second, results cannot necessarily be generalized to the whole study population, but hold true only for individuals near the cutoff.

In the end, SPARK decided to postpone the impact evaluation to future years of the project. After running power calculations, sample sizes were too low in this cohort and the training and coaching content were still undergoing some changes.

M&E results

- **Participation:** In 2012, Birzeit recorded 713 graduates trained in the program, 56 percent of whom were women. In that same year, six business plan competitions were held. In total, 194 business plans were submitted, 44 percent of them by women. Overall, 20 percent of program participants submitted a business plan.
- **Matching to finance:** In 2012, 18 new financial match-ups were finalized. This shows that only a small fraction of the business plans materializes in financial match-ups.
- **Start-ups:** Counting both years 2011 and 2012, five business start-ups succeeded out of the 18 financial matches made (28 percent), and three businesses are considered grown.⁵ This number is likely to increase once the data collection results from 2013 become available.

Lessons learned

Pilot-testing instruments is crucial

Pilot-testing is a crucial step in the implementation of a data collection tool. For instance, the length of the questionnaire was reduced significantly when it emerged that respondents lost interest after 30 minutes. Further testing may lead to additional adjustments, trying to find a compromise between reducing the time burden for respondents and keeping the quality and the usefulness of the information collected.

Pilot-testing has also proven to be extremely valuable when using Salesforce. Indeed, when the data collection officers first entered real data from the M&E tools into Salesforce, some practical issues arose, such as misalignments between the available data and the template to be filled.

Allocate sufficient staff time for M&E activities

Throughout the Taqueem journey, SPARK and Birzeit have acknowledged having underestimated the time required performing the designed M&E activities, despite having a person dedicated to M&E in their headquarters. From their own words, what was lacking was a dedicated local staff member able to focus “exclusively on the [M&E] project.”

Activities are also very time-consuming when numerous actors are involved. For example, SPARK faced challenges in getting their key local partners in Palestine to start gathering and inputting data. They recommend other practitioners to prioritize showing the partners the value of supporting this form of data collection.

⁵ A business is considered grown when it has been started in the previous outcome reporting period (half-year) and has shown turnover growth in the next; outcome business survival rate is a separate indicator (three years of turnover shown).

Dealing with Salesforce

The implementation of the Salesforce data management software caused some delays to SPARK and Birzeit. In order to successfully adapt an M&E system to Salesforce, an organization needs to nail down, at the detailed level, what information they want to collect as part of their M&E process. This not only refers to the indicators, but also the question fields, and the pick list values they will have as options in each field.

Next steps

Finalize the implementation

The finalization of the implementation of the designed M&E system, including a functioning data management system, remains a high priority for SPARK and Birzeit. Moreover, additional training on data entry using Salesforce is expected to increase the effectiveness of output monitoring.

Replicate the data management system

Provided the implementation of the system is successful, SPARK seeks to replicate the M&E approach to similar youth entrepreneurship programs in other supported countries and organizations. Given the experience gained with Birzeit, the replication promises to be less time-consuming than the initial pilot of the system.

Programmatic adjustments

The data already collected indicate that only a very small number of program participants end up starting a business through the financial match-up made by the university. Thus, SPARK and Birzeit have initiated several programmatic adjustments: (i) monitor trained graduates more closely after they leave the Birzeit program to determine SME start-up and job creation effects; (ii) increase the visibility and attractiveness of the Spark guarantee loan system with Faten, the Palestinian MFI providing start-up loans, so that more graduates access these; and (iii) develop higher-level coaching services for new entrepreneurs that do not exist on the market.



Making Learning Count

Effective Monitoring and Evaluation of Youth
Employment Programs in the Arab World

In response to the lack of evidence behind youth employment interventions in the Arab world, the ILO joined with partners, including Silatech, the World Bank, the Jacobs Foundation and the Swedish International Development Foundation in 2009 to establish the Taqeem Initiative. Taqeem supports youth employment policy-makers and practitioners in the Arab world in enhancing the monitoring and evaluation of their programs.

“Making Learning Count” is a collection of case studies telling the story of the evolution of M&E systems in seven youth serving organizations. The volume seeks to contribute to the knowledge and learning around “what works” in youth employment and how organizations can contribute to this knowledge base and to provide important insight on how the knowledge base can be expanded by strengthening learning and accountability across organizations and programs, large and small.

International Labour Office
Youth Employment Programme
Employment Policy Department
Route des Morillons 4
CH-1211 Geneva 22

youth@ilo.org