THE Rockefeller Foundation

EVALUATION OFFICE

DISCUSSION PAPER

Emerging Opportunities: Monitoring and Evaluation in a Tech-Enabled World

Linda Raftree and Michael Bamberger

September 2014

itad

Financial support provided by
THE ROCKEFELLER FOUNDATION

THE ROCKEFELLER FOUNDATION

About Itad

Through its innovative consultancy services in monitoring and evaluation, Itad provides the insight and ideas to ensure resources invested in international development deliver the best possible results for the poor. Established in 1984, Itad focuses on making international development as effective as possible by giving organizations the information and insight they need to make development work smarter and produce better results. Itad thrives on complex, meticulous monitoring and evaluation work involving many different stakeholders across many different scenarios. It's special areas of work include evaluation and impact assessment, results-based M&E systems, capacity development and training, project management services, and research and advisory services.

About the Rockefeller Foundation Evaluation Office

For more than 100 years, the Rockefeller Foundation's mission has been to promote the well-being of humanity throughout the world. Today, the Rockefeller Foundation pursues this mission through dual goals: advancing inclusive economies that expand opportunities for more broadly shared prosperity, and building resilience by helping people, communities and institutions prepare for, withstand and emerge stronger from acute shocks and chronic stresses. Committed to supporting learning, accountability and performance improvements, the Evaluation Office of the Rockefeller Foundation works with staff, grantees and partners to strengthen evaluation practice and to support innovative approaches to monitoring, evaluation and learning.

THE Rockefeller Foundation

EVALUATION OFFICE

Financial support for this paper was provided through a grant from the Rockefeller Foundation to Itad.

Cover photo: The Rockefeller Foundation

Table of contents

LIST OF BOXES

LIST OF TABLES

LIST OF FIGURES

Acronyms

CAPI	Computer-assisted personal interviewing
CEASE	Center to End All Sexual Exploitation
CLEAR	Center for Learning on Evaluation Results (South Asia)
CRS	Catholic Relief Services
EPPI-Centre	Evidence for Policy and Practice Information Co-ordinating Centre
GPS	Global positioning system
HARITA	Horn of Africa Risk Transfer for Adaptation
HHI	Harvard Humanitarian Initiative
ICT	Information and communication technology
ICTME	ICT in monitoring and evaluation
ICT4D	ICT for development
ITU	International Telecommunications Union
MIT	Massachusetts Institute of Technology
M&E	Monitoring and evaluation
MDG	Millennium Development Goals
MIS	Management information system
MSC	Most significant change
ODI	Overseas Development Institute
PDA	Personal digital assistant
PRA	Participatory rural appraisal
RCT	Randomized control trial
RIWI	Real-time Interactive Worldwide Intelligence
SMS	Short message service
SPEED	Smart Power for Environmentally-sound Economic Development
TOC	Theory of change
UNDP	United Nations Development Programme
URL	Uniform resource locator (Internet)
WHO	World Health Organization (UN)

About the authors

Linda Raftree has worked at the intersection of community development, participatory media and information and communication technologies (ICT) since 1994. She advises the Rockefeller Foundation's Evaluation Office on the use of ICTs in monitoring and evaluation. Linda has also worked with Plan International USA on innovation, transparency and strategy. She has conducted research on adolescent girls and ICTs for UNICEF, the role of ICTs in child/youth migration for the Oak Foundation, and the use of mobile technologies in youth workforce development for the mEducation Alliance. Linda coordinates Technology Salons in New York City and she advocates for greater dialogue and discussion around the ethics of ICT use and data privacy in the humanitarian and development space. She writes 'Wait... What?' a blog about new technology and community development, and tweets at @meowtree.

Michael Bamberger has evaluated development programs in more than 40 countries in Africa, Asia, Latin America, and the Middle East. He has worked with several NGOs in Latin America and with the World Bank in the areas of development evaluation, capacity development and gender. He has also consulted on development evaluation with a number of UN, multilateral, and bilateral development agencies; published widely on development evaluation; and served on the editorial board of several evaluation journals. He advises the Rockefeller Foundation's Evaluation Office, and he was recently appointed to the UNDP Evaluation Office International Evaluation Advisory Committee. He is currently co-editing and co-authoring "Addressing complexity in development evaluation: a practitioners guide" (scheduled for publication in September 2015) which includes several chapters on the contribution of ICT to the evaluation of complex programs.

Acknowledgements

We would like to express our heartfelt gratitude to everyone who contributed to the compilation of this report over the past year. This includes those who shared their ideas, efforts and challenges in Technology Salons and other similar discussion sessions in New York City, Washington DC, Nepal, Ghana and Cameroon, and those who provided further information about their work and ideas through personal discussions and email.

We especially thank our group of expert reviewers for their feedback and advice: Patricia Rogers, Professor of Public Sector Evaluation at RMIT and lead at Better Evaluation; Paula Lytle, Senior Social Development Specialist, World Bank; Emmanuel Letouzé, Co-founder & Director of the Data-Pop Alliance on Big Data and Development; Laura Walker Hudson, CEO, Social Impact Lab Foundation; and Kerry Bruce, Senior Director, Global Health and Measurement, Pact.

Last but not least, we thank Nancy MacPherson and Jill Hannon of the Rockefeller Foundation Evaluation Office for their constant support and guidance. Without the Rockefeller Foundation's dedication to exploring the potential as well as the challenges of new technology tools and related methods in the field of monitoring and evaluation, this report would not have been possible.

Linda Raftree and Michael Bamberger

© George Henton

Foreword

Inspired by the potential for technology-enabled tools to contribute to the evolution of the monitoring and evaluation (M&E) field, and by the information and communication technology (ICT) innovators we have met in the course of our work at the Rockefeller Foundation, we are pleased to provide financial support for this paper as part of a wider effort to promote innovation in evaluation.

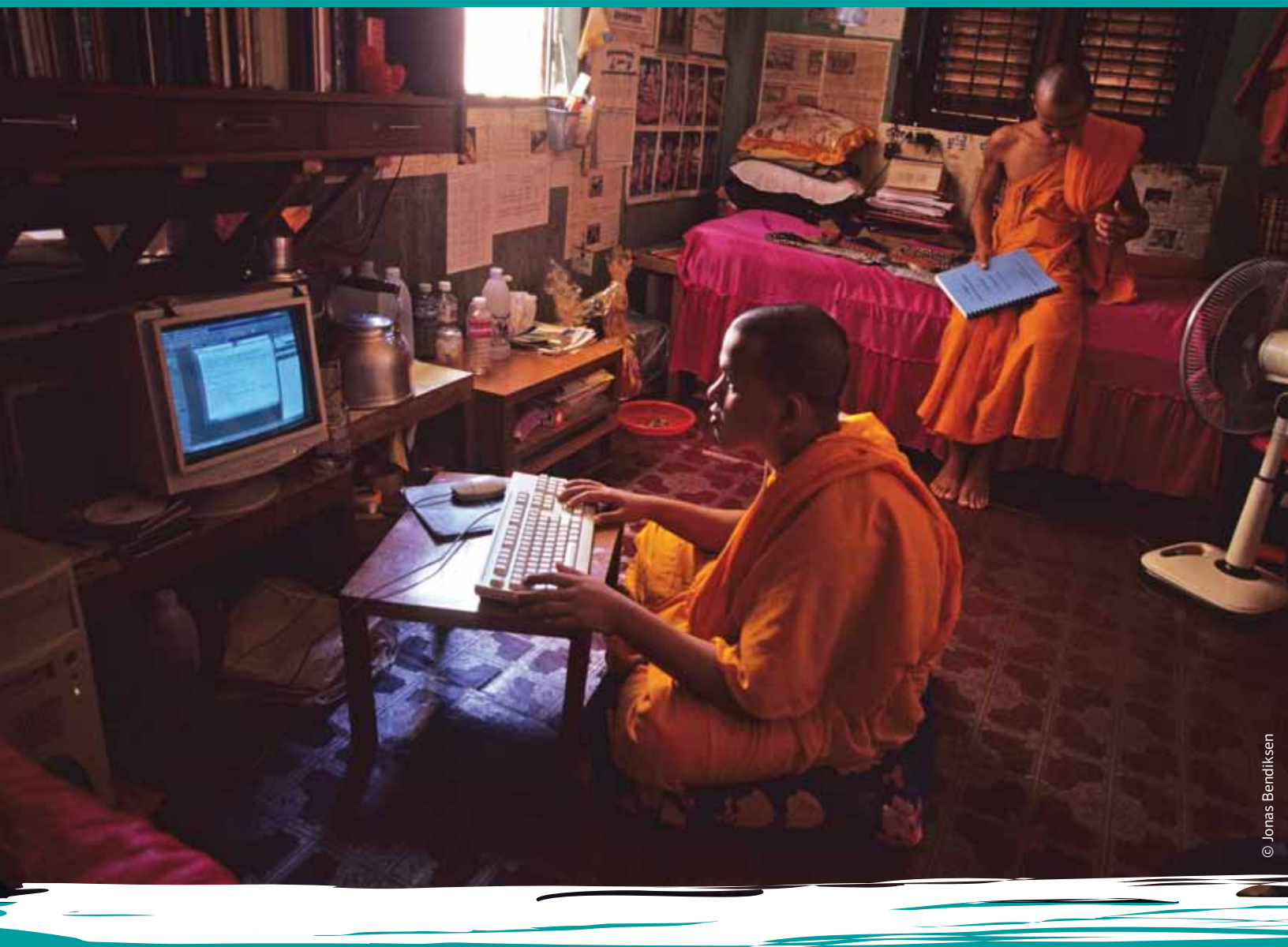
The increasing complexity of development coupled with the widening range of public, not-for-profit and private sector actors and the demand for more timely feedback has challenged the utility of conventional approaches to monitoring and evaluation in many development contexts. Though emerging information and communication technologies offer the promise of including more voices in a more timely way than conventional methods, the methodological rigor of technology-enabled M&E has sometimes been questioned and viewed as unreliable in contemporary evaluation debates, Despite great strides in the rapid adoption and proliferation of technology throughout the world, evaluation practice has remained largely paper-based. As a result, traditional evaluation methods and approaches to learning, accountability and feedback have often not kept pace with the significant advances in technology.

In spite of this broad reluctance, M&E innovators are already experimenting in this new space and harnessing the power of technology to confront both real-world evaluation constraints and fundamental methodological challenges. By reflecting on ways in which these innovators have begun to navigate new territory, and by exploring the great potential for technology to further transform and advance traditional evaluation methods, this paper aims to highlight the current state of tech-enabled M&E while also maintaining a critical perspective which recognizes the limitations and inherent risks which evaluators should remain mindful of when engaging in this new and exciting space.

In this paper, the authors highlight some of the ways that ICTs are helping overcome common M&E challenges, including "real-world" challenges and methodological and conceptual challenges. The paper also offers ideas on untested areas where ICTs could play a role in evaluation, and an in-depth discussion of some of the new challenges, problems and risks that arise when incorporating ICTs into the M&E process as a whole. Finally, it offers a checklist for thinking through the incorporation of ICTs into M&E.

As we continue to explore and apply new technology in our work at the Rockefeller Foundation and to learn from M&E innovators, we hope that this initial landscaping of ICTs in M&E serves as a launching point for further discussion, learning and improved M&E practice, all in the service of better development outcomes for humanity.

Nancy MacPherson
Managing Director, Evaluation
The Rockefeller Foundation



© Jonas Bendiksen

Executive Summary

Background

Various trends are impacting on the field of monitoring and evaluation in the area of international development. Resources have become ever more scarce while expectations for what development assistance should achieve are growing. The search for more efficient systems to measure impact is on. Country governments are also working to improve their own capacities for evaluation, and demand is rising from national and community-based organizations for meaningful participation in the evaluation process as well as for greater voice and more accountability from both aid and development agencies and government.

These factors, in addition to greater competition for limited resources in the area of international development, are pushing donors, program participants and evaluators themselves to seek more rigorous – and at the same time flexible – systems to monitor and evaluate development and humanitarian interventions.

However, many current approaches to M&E are unable to address the changing structure of development assistance and the increasingly complex environment in which it operates. Operational challenges (for example, limited time, insufficient resources and poor data quality) as well as methodological challenges that impact on the quality and timeliness of evaluation exercises have yet to be fully overcome.

A second trend, happening in parallel to these changes in the international development and evaluation space, is the explosive growth of mobile phones and other information and communication technologies (ICTs) at all levels of society around the globe. Greater access to digital devices, especially the mobile phone, is changing how people access information, how they communicate, and how they engage with services and each other. Increasing attention to and sophistication of digital tools is permeating the sphere of development as well. New tools and approaches are rapidly making their way into the area of M&E, yet many M&E practitioners have not explored their potential.

The current paper offers a broad overview of how ICTs and digital tools are being used to help bring M&E up to speed with the changing external environment and ways that they are helping to address operational and methodological challenges. Based on an examination of the available literature; in-depth discussions with development, technology and evaluation practitioners; and interviews with development and evaluation experts from a range of disciplines who are working to find new ways to measure progress and impact qualitatively and quantitatively and to learn and improve practices, the paper offers a snapshot of a wide range of ways that ICTs are being integrated into M&E.

Key Messages

1. **ICTs are being used throughout the planning, monitoring and evaluation cycle, but there is little hard evidence of their effectiveness.**

   There is quite a bit of experimental use of ICTs in M&E, yet much of it is not well-documented in terms of its usefulness in overcoming operational and methodological challenges or improving the M&E process. It can be difficult to convince donors or management that using a new tool is a useful investment due to the lack of evidence.

   Some examples of how experimentation is happening in the various stages of the M&E cycle include:

   - **Diagnosis.** ICTs are being used to bring new voices and broader participation into program diagnosis and enable a wider range of input at a reduced cost. They are enabling evaluators to better manage and pull possible trends out of large data sets.
   - **Planning.** ICTs are being used to help achieve greater inclusion in planning processes. New technologies make it easier to compare and visualize data sets and to analyze data based on location so that resources can be better allocated. Data are also being aggregated more quickly and shared at various levels to improve participation in the planning process and to make better decisions. New software tools are being used to enhance the development and management of theories of change.
   - **Implementation and monitoring.** ICTs are allowing for the collection of real-time data on participant experiences, behaviors and attitudes, meaning that analysis can be conducted early in the process and course corrections can be made to improve interventions and outcomes. Direct feedback from program participants is also being made possible through new ICTs, and it is assumed that this can help achieve greater transparency and accountability.
   - **Evaluation.** ICTs can be integrated to increase the voice of vulnerable and underrepresented groups and broaden the types and volume of data being collected, combined, compared and analyzed. New technologies may be able to help overcome challenges and constraints such as sample bias and poor data quality, and improve the understanding of complex sets of behavior and data.
   - **Reporting, sharing and learning.** ICTs are enabling wider circulation of evaluative learning, interactive sharing, and greater public engagement with evaluation findings.

2. **A number of areas with potential have not been sufficiently explored.**

   It is supposed that ICTs could play a role in improving the validity of methods such as sample selection strategies (for example random routes), improving baselines or reconstructing baselines, reducing sample bias, enhancing rating scales, supporting concept mapping, evaluating complex interventions and improving qualitative case study methods. Many new ideas have not been tested and explored, however, and further work could be done to experiment with some of the ways that ICTs could support these methods.

3. **ICTs bring new challenges that evaluators need to prepare for and address.**

   Some of these new challenges include:
   - **potential for selectivity bias** when those who do not have access to or strong capacity to use ICTs are left out
   - **potential for tool- or technology-driven M&E processes** when M&E plans are adapted to ICT tools rather than ICT tools being selected because they can help meet the needs of an M&E plan
   - **overreliance on digital tools, data and numerical indicators,** which may lead to a loss of quality control measures, over-collection of data with little capacity to analyze it or provide context, and the loss of the personal rapport and contextual understanding obtained from project visits and face-to-face interviews when these are replaced with rapid and often remote electronic data collection
   - **low institutional capacity and resistance to change,** which are common challenges for organizations that do not have the budget to fully train and integrate ICTs into their operations
   - **loss of privacy and increased levels of risk for evaluation participants,** which can result if data and privacy are not carefully protected, and if a thorough risk assessment is not conducted to plan for potential negative or unintended consequences.

4. **Careful planning and analysis around the use of ICTs can help mitigate risk and improve outcomes.**

   A number of lessons were drawn from conversations with practitioners and a review of the literature as well as blog posts and other less formal documentation. These lessons, which can serve as a starting point for those wishing to begin integrating ICTs into their M&E practice, include:
   - **develop a quality M&E plan** before thinking about how ICTs can help with implementation
   - **address threats to evaluation validity** including internal design validity, statistical conclusion validity, construct validity, and external validity
   - **determine whether and how ICTs can add value to an M&E** plan rather than forcing ICTs into a plan or starting with ICTs and creating a plan that revolves around them
   - **select or assemble the right combination of ICT and M&E tools** because no one ICT or M&E tool is likely to offer everything that an organization or evaluator is looking for
   - **adapt and test the tools and the process with different audiences and stakeholders** and iterate them along the way to be sure they are appropriate and user friendly and that they work for different stakeholders
   - **be aware of differing levels of access and inclusion** because marginalized members of a community or group may be left out if ICT-enabled M&E is not designed with inclusion in mind
   - **understand people's motivation** to participate in M&E – including program participants, staff, government, and any other stakeholders – and designing accordingly
   - **ensure privacy and protection** so that the M&E process is not putting people at risk
   - **identify potential unintended consequences** that could result from the introduction or use of ICTs, including domestic violence against women, theft, harassment from authorities,

competition for devices, or misinterpretation of the goals of the evaluation exercise

- **build local capacity and base processes on local systems** rather than bringing in top-down, externally created solutions that may be costly and unsustainable in the longer term
- **measure what matters** by ensuring that tools are not leading the process, but rather they are enhancing the collection of the data needed for the evaluation process
- **share M&E information effectively with program participants and staff** and use it to support better decision-making, learning and improvement.

Conclusion

The field of ICTs in M&E is emerging and there is activity happening at multiple levels and with a wide range of tools and approaches. The field would benefit from a greater effort at documentation, as evidence of the utility and impact of ICTs for M&E is still underdeveloped. Increased pressure to show impact may open up space for testing new approaches, and some of those highlighted in this paper could serve as a starting point for exploration. At the same time, a number of pitfalls have been signaled and these need to be considered when designing an evaluation plan that involves ICTs. The checklist offered in this paper can help evaluators think through known challenges and identify other barriers and potential risks. An investment in the development, application and evaluation of innovative new M&E methods could help evaluators and organizations adapt their approaches throughout the entire program cycle, making them more flexible and adjusted to the complex environments in which development initiatives and M&E take place.

© Antony Njuguna

© Linda Raftree

1

Introduction

Greater competition for the limited resources available for international development assistance, combined with the broadening expectations of what development assistance should achieve, has heightened the demand for efficient systems to assess the performance and impact of international development programs. Developing country governments are also increasing their commitment to building systems that can assess the performance of national development plans, as evidenced by a steady growth in the number of developing countries that are implementing national evaluation policies. In addition, as civil society and local organizations gain greater voice, there is a heightened demand to assess the participatory, humanitarian and equity-focused dimensions of development and to include program participants more meaningfully in monitoring and evaluation (M&E) processes. Finally, the growing scope of human-made and natural crises has increased the demand for assessing the impacts of development during crises and in unstable environments.

All of these factors are creating a greater demand for more rigorous – and at the same time more flexible – systems to monitor and evaluate development and humanitarian interventions. Critical assessment of the strengths and limitations of current approaches to M&E has identified serious limitations of many existing approaches for addressing the changing structure of development assistance and the increasingly complex environment in which it operates.

Alongside this wider context, there are two broad sets of historical challenges in conventional M&E approaches. The first set is often referred to as "real-world" or operational challenges, while the second set can be categorized as methodological challenges. Emergent ICT tools and applications may have potential to help address some of the overarching M&E challenges in the wider development space while, at the same time, contributing to overcoming real-world and methodological challenges.

1.1 Real-world challenges

Generally, evaluations and monitoring systems are conducted and created under "real-world" constraints, meaning they operate within a limited budget and have limited access to important data. In addition, they must be designed, implemented, analyzed and disseminated under severe time con-

straints.[1] The following lists some common challenges resulting from real-world constraints.

- The cost of collecting the desired data is too high to be feasible within the approved budget. This means that the sample size has to be reduced or that important qualitative data collection methods which complement the quantitative survey, such as focus groups or in-depth interviews, cannot be included in the evaluation design.
- The range and complexity of potentially relevant data is expanding exponentially through the availability of big data,[2] and the speed and ease of collecting new kinds of data. Most current M&E systems do not have the technical expertise and logistical or financial resources to capture, or utilize all of this data.
- It is difficult and expensive to obtain up-to-date information on how target populations use the services offered by government or development institutions.
- Certain groups are difficult or expensive to reach, such as drug users, sex workers, ethnic minorities, poor households in remote areas or families who have moved from their original addresses. Consequently, they are often excluded or underrepresented in the evaluation.
- When evaluations are conducted in conflict zones or in dangerous communities, it is difficult for interviewers to reach some areas or it is risky for respondents to be seen talking to interviewers. In other cases, cultural constraints make it difficult for women to travel outside their compounds or meet with interviewers.
- Observing how a project is implemented can be just as important as measuring the changes that have taken place over the life of the project (the conventional pre-test-post-test comparison). However, observing the implementation process is time-consuming, expensive and complicated, and is often excluded from the evaluation.
- Development programs are affected by the economic, political, socio-cultural, demographic and ecological/environmental contexts within which they operate. However, due to the costs and complexity of collecting information on these contextual factors, they are often not systematically incorporated into the evaluation design.
- Finally, validity of the evaluation findings depends in large part on the quality of the data collected. Unfortunately, data quality control is such an expensive and time-consuming process that, when working under budget and time constraints, it may not be possible to follow standard quality control procedures, such as checking to ensure the right subjects have been selected and interviewed, or ensuring that questions are asked correctly, in the right order and with the correct follow-ups.

1.2 Methodological challenges

In addition to real world constraints, methodological challenges impact on the quality and timeliness of evaluation exercises. The list below identifies some examples of these challenges.

- Most widely-used M&E systems were designed to study relatively stable programs, with a well-defined set of outcomes that were expected to be achieved through fairly simple causal paths. However, today, many development agencies are moving towards multi-donor programs, with

[1] There are two kinds of time constraints. The first refers to the time period during which the evaluation must be conducted. For example, the terms may stipulate that the evaluation must be designed, implemented, analyzed and the report presented within a certain number of weeks, and the evaluator may feel this does not allow sufficient time. The second refers to the stage of the project at which the evaluation is conducted. For administrative reasons, many evaluations that are intended to assess project impacts are conducted at a point in the project when it is still too early to assess impacts.

[2] "Big data" is an umbrella term that refers to one or more of three trends: i) the volume of digital data generated daily as a by-product of people's use of digital devices is growing; ii) new technology, tools and methods are available to analyze large data sets that were not specifically created for the purpose of analytics; and iii) policy-making insights are being extracted from these data and tools. Section 3 of this paper expands on the role of big data in M&E.

multiple and often changing interventions, and responsibility for the programs and management of resources is increasingly moving to host-country agencies. Thus, it is recognized that conventional evaluation designs that work for "simple" or "complicated" programs no longer work for the emerging complex programs that must be implemented in fast-changing environments that face factors such as climate change, ongoing conflict, and weak or failing states. This new focus on complexity also recognizes the interconnectedness of programs being evaluated with other features of the country or the international environment. Complexity is a challenge that the evaluation community is only starting to address (Box 1 provides a brief definition of the concepts of "simple", "complicated" and "complex" programs.)

- Programs intended to produce behavioral change, such as reducing high-risk sexual

BOX 1

Defining simple projects, complicated programs and complex interventions

Simple projects:

- include relatively simple "blue-print" designs that produce a standardized product
- follow a causal path that is relatively linear
- have defined start and end dates making it time-bound
- have only a few objectives but they are clearly defined
- define a target population that is usually relatively small
- have a well defined budget and resources.

Complicated programs:

- include a number of different projects each with its own "blueprint"
- follow causal paths for different components and different objectives, but are still relatively linear
- have information on the process of project implementation that is often not well documented
- target a larger and more diverse population
- involve several different donors and national agencies
- may be implemented by different donors in slightly different ways
- set objectives in broader and less clearly defined terms
- set up without start-end dates, thus not so time-bound
- focus on the importance of program context
- merge funds into ministry budgets, making it difficult to estimate.

Complex interventions:

- merge into national or sector development policy, making specific program interventions difficult to identify
- follow non-linear causal paths, as there may be multiple paths to achieve an outcome, or the same set of inputs may produce different outcomes in different settings
- are delivered by multiple agencies, and components and service are not delivered in a uniform manner
- have emergent designs that evolve over time
- have program objectives that are difficult to define or not specified
- have non-proportional relationships between inputs and outcomes.

SOURCE: Adapted from Bamberger *et al.*, 2012.

NOTE: the listed items are examples of how these are often set up. Not all projects will contain all of these elements.

behavior or improving public service agencies, often interact with vulnerable and socially marginalized groups. In other cases, behavioral change processes influence intended outcomes that the program cannot predict or control, such as how interactions among members of the target population affect program implementation. As these responses are often unintended, it is difficult to track or even to identify them with conventional evaluation data collection methods.

- Theories of change (ToCs) are increasingly used in evaluation design. These are difficult to develop and update in a participatory way when there are multiple stakeholders based in widely dispersed geographical locations. Consequently many TOCs become rapidly outdated or do not have mechanisms to incorporate new information or changing contexts. TOCs make it possible to constantly test and revise the assumptions built into the model and the assumed linkages between different levels of the model. Yet most evaluations do not have the capacity to constantly update the TOC and, as they become outdated, they fail to make their potential contribution to the implementation and interpretation of the evaluation.
- Quantitatively oriented evaluations (e.g. randomized control trials and survey-based designs in general) find it difficult to collect qualitative data such as leadership styles and patterns of interaction among household members. The attempt to measure these complicated or complex[3] multi-dimensional phenomena through a small number of simple quantitative indicators can result in the problem of construct validity.[4]
- Programs may operate in insecure locations such as militarized areas and places with high crime or gang activity. These sites may be dangerous for evaluators to visit. Other areas may be geographically isolated or otherwise difficult to reach, making it too expensive to capture information about them using conventional data collection tools.
- Most programs are affected by broader contextual changes such as population movements, climate change, the condition of transport networks and soil erosion. These are also difficult to capture through traditional evaluation methods.

Evaluators and practitioners are experimenting with ICTs to include the voices of participants and beneficiaries of development programs ...

- The focus of development is shifting towards complex, multi-component, multi-agency programs with a range of difficult-to-document interventions that can reach into the hundreds. These programs also have a wide range of outcomes that are often not clearly defined. With conventional evaluation designs unable to assess the complex interventions, the evaluation community is searching for new methodologies for evaluating complex programs.
- Evaluation designs continue to struggle with the challenges of: i) internal validity, or reasons why an inference about a causal relationship between two variables – e.g. a project intervention and an observed outcome – may not be

[3] The evaluation literature distinguishes between simple, complicated and complex interventions and simple, complicated or complex evaluations. A complex intervention is characterized as having multiple components that can produce multiple outcomes through multiple, and usually non-linear causal pathways, where relations between causes and effects (inputs and outcomes) are non-proportional (small changes in inputs can produce large changes in outcomes and vice versa). Outcomes may not be known in advance and the program design may be emergent (i.e. it evolves and does not always follow a predictable path). In contrast, a complicated program or intervention has multiple partners, multiple components that are often not implemented in a standard way, not clearly defined, and do not have uniform implementation procedures (each agency may follow a different path). However, the causal pathways are relatively linear.

[4] "Construct validity" refers to the different reasons why the constructs used to measure inputs, processes, outcomes and impacts may not be appropriate.

Table 1: Potential applications of ICTs to address common real-world budget, time and data challenges

CHALLENGE	PROMISING ICT APPROACHES
Data collection costs are high	Collecting and analyzing survey data through smart phones and hand-held devices can eliminate costs of printing and transporting survey instruments, and dramatically reduce costs of data analysis. Managing M&E processes and enumerators using software can cut costs and improve efficiency. Collecting mobile data can enable errors to be caught at the point of contact and lower the need to return to re-collect data. Collecting digital data eliminates the need for double data entry.
Real-time information on service use by the target population is hard to obtain	Smart phones can monitor whether clients follow-up on automated phone messages. Smart phones allow for review of application data to understand a "user's journey" through the application and how he or she is using the application. Users of a service can provide input directly via SMS when it is most convenient for them, which may be more convenient than finding time to join focus groups or be interviewed and surveyed according to others' schedules and time frames.
Some groups are difficult to reach	SMS-based surveys can be used to reach out and collect data. SMS is one of the most wide-spread "lowest common denominator" technologies available.
Some groups are dangerous to reach and interview	SMS-based surveys and self-reporting tools via the web can reduce risk to evaluators and those they are interviewing. Phone interviews can be done in areas that are insecure. Incident reporting via phone and Internet allows for more widespread input and self-reporting.
The process of project implementation is difficult to monitor	Smart phones can record video and audio during project implementation activities such as meetings, work groups or classroom activities. Web-based M&E platforms allow for better documentation of processes as well as outputs and outcomes.
Data collection on contextual factors that affect program outcomes is difficult and expensive	Smart phones and Internet enable integrated access to secondary data sources. Big data provides access to more extensive contextual data.
Quality control of data collection is expensive to ensure	GPS-enabled devices can check that the interviewer is in the correct location. Electronic versions of surveys can ensure that questions are asked in the correct order and can include automatic consistency checks. Audio recording can be randomly activated so that the supervisor can listen to the interview. Video can record body language used during an interview or survey, which enables evaluators to understand more. Hand-held devices provide real-time feedback so that errors can be identified and corrected before the interviewer/enumerator leaves the site.
Behavioral change needs monitoring	Video and audio recordings at project locations, in the community or in households improve capacity to monitor behavior directly. Socio-metric analysis of patterns of interaction and communication can be conducted in the community or organization.

Table 2: Potential applications of ICTs to address common methodological challenges

CHALLENGE	PROMISING ICT APPROACHES
Theory of change is needed for multiple, dispersed stakeholders	Online theory of change software permits people in different locations to participate in the design and updating of the TOC. Some tools automatically change text boxes into photos when working with communities with low literacy.
Quantitative evaluation designs need to incorporate qualitative data	Online software permits video and audio data captured on cell phones to be coded. Online software can allow respondents to classify statements and concepts using their own criteria.
Geographic dimensions of programs need capturing	GPS-enabled devices can be used to construct maps locating events, services and important features of the community or area.
Broader contextual factors affecting program outcomes need capturing	Satellite images can track physical change over large areas, such as population movements, rainfall patterns, location and size of settlements, effects of climate change, and location and quality of infrastructure. Crowdsourcing can provide real-time feedback on damage from natural disasters, ongoing and planned political protests and outbreaks of disease.
Complex programs* require development of specifically targeted evaluation applications	Applications can integrate multiple data monitoring sources of social media communications, enabling the study of attitudinal and behavioral change. ICT can allow for modeling of complex systems and causal pathways. ICT can assist in the development and configuration of case study analysis. Software permits the development of scales and indices (such as concept mapping) that define and rate complex concepts.
Data silos need to be reduced	ICTs can help move organizations towards common data definitions (e.g. numbering systems for regions, districts, health centers, water points, communities and commonly defined indicators across programs. Use of a management information system can eliminate the need to re-collect data over and over again.

* Some of these potential applications are still work in progress and have not yet been widely tested or documented.

valid, and ii) external validity, or reasons why inferences about how evaluation findings would hold in other settings may not be valid. Well designed quantitative evaluations can usually address the internal validity issue, but they have difficulty in addressing external validity. Thus, there is an increasing use of mixed-method designs because their careful combination of quantitative and qualitative methods can provide more reliable estimates of both internal and external validity.

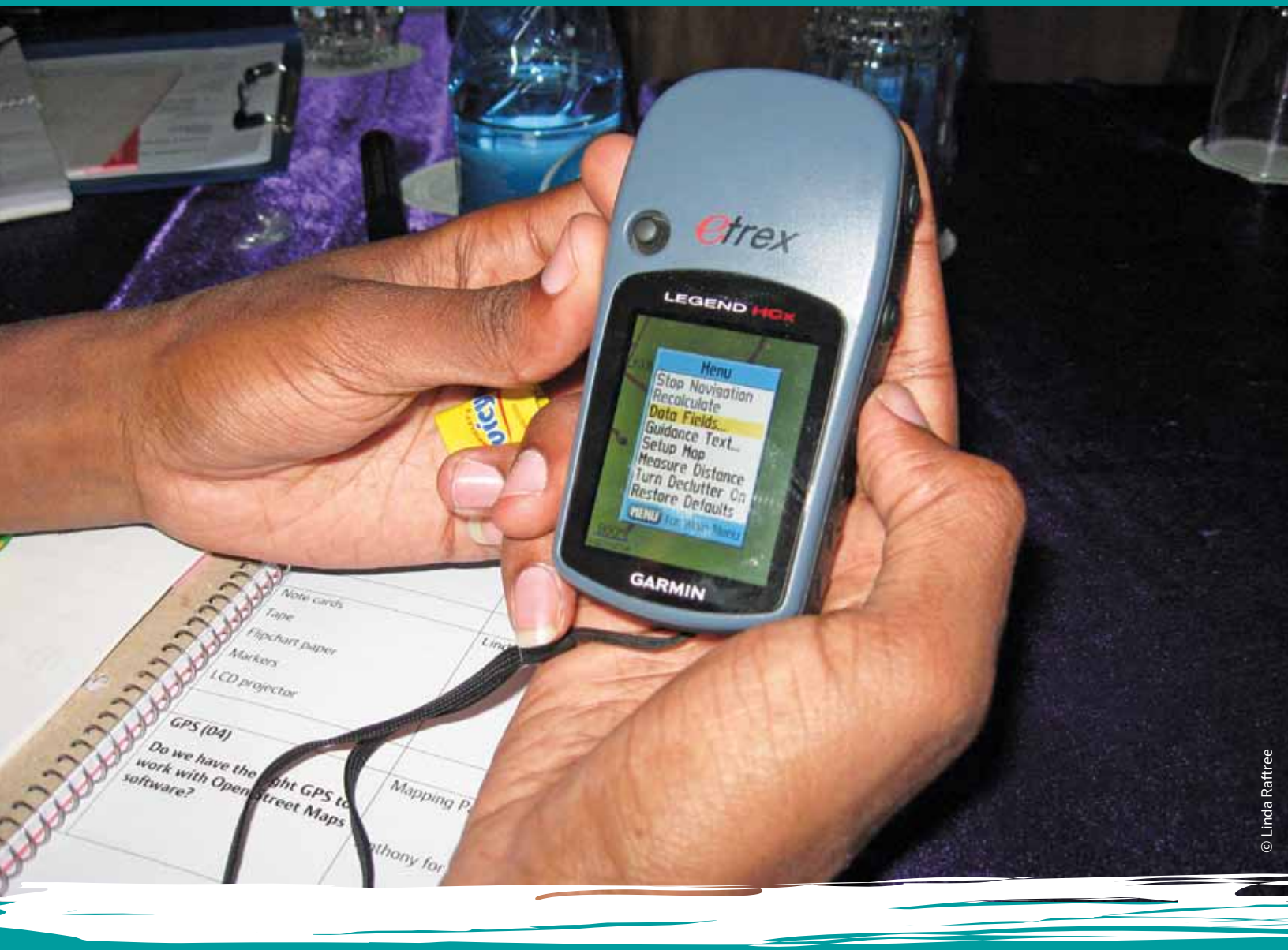
ICTs are being used to help bring M&E up to speed with the changing external environment and to address some of the real-world and methodological challenges mentioned above. Evaluators and practitioners are experimenting with ICTs to include the voices of participants and beneficiaries of development programs, allowing them to weigh in on what success should look like and make possible a more realistic evaluation of whether or not success has been achieved. Evaluation teams are using ICTs to help improve efficiency and quality of data, and to reduce sample bias. They do this by providing access to better data to construct the sample frame, reaching vulnerable and difficult-to-reach groups that are frequently under-represented, and improving quality control of the interview process.

Large data sets and improved data processing capacities are allowing researchers and evaluators to identify formerly unseen patterns that require further investigation. ICTs are also playing a role in enabling wider sharing and discussion of evaluative knowledge which, in turn, helps development practitioners avoid repeating mistakes and failures. It also allows dissemination of evaluative knowledge to a wide audience, outside of boardrooms and program teams, in order to stimulate broad discussion and learning. New technologies are also being used to facilitate training of developing country evaluators, helping to build capacity and knowledge that will enable local evaluators and institutions to play stronger roles in the evaluation process in their own countries (Rodin and MacPherson, 2012).

Much of the attention around ICTs in M&E focuses on enhancing the participation of program participants in feedback loops that seek to improve transparency and accountability in aid and development or government service programs. ICTs are being used to increase voice and participation throughout the program cycle – from diagnosis, through planning and implementation, to evaluation and the dissemination of evaluative knowledge. Gathering a wider perspective from a broad network, learning from experimentation through results testing, setting up and learning from lessons and feedback loops, and having the ability to capture the value of both successes and failures have been identified as key elements of organizations with strong capacity to innovate. ICTs can play a role in facilitating these capacities within organizations (The Rockefeller Foundation, nd). This is especially important as development programs and their accompanying evaluations are increasingly understood to be complex systems.


© Josh Nesbit

Table 1 lists some of the promising approaches that are discussed in this paper. While some of the approaches are already well documented, others are included as new areas to explore.



© Linda Raftree

2

Information and communication technologies in the M&E cycle

New ICTs impact virtually every aspect of people's lives across the globe. A 40 percent rise in mobile broadband subscriptions was seen at the global level in 2011, access to and use of affordable tablets and other devices is growing steadily, and the International Telecommunications Union (ITU) reported that growth in ICT uptake in 2012 was almost universal. Mobile cellular subscriptions had reached almost 7 billion by the end of 2013, with mobile network coverage expanding to more and more remote areas (ITU, 2013a; 2013b). Those working in international development are devising a myriad of ways to take advantage of this growth in ICT access and use. Incorporation of ICTs into development work, a field known as ICT for Development (ICT4D), is expanding and changing at the same rapid pace as technology itself.

ICTs came to the forefront in the 1960s, when the public sector began using information technology systems to support administrative functioning. In the 1980s, multinational corporations began seeing computers as tools that could deliver economic growth in the private sector. In the 1990s, which saw the uptake of the Internet and launch of the Millennium Development Goals (MDGs), people began thinking about how ICTs might be used for development efforts. By the year 2000, the integration of ICTs into development programs had become commonplace. Into this environment arrived the cellular phone, offering unprecedented opportunity because of its widespread use and adoption even in poor communities. Its rapid uptake around the world renewed emphasis on ICT4D.

> *Increasingly, ICTs are enabling improved feedback and participation from the populations that development agencies serve.*

As development theories have advanced, the field of ICT4D has also moved forward. Today, ICT4D often places emphasis on participation, improvisation, flexibility, learning and local capacity. The successful ICT4D initiatives are not developed for the poor in a laboratory. Rather, they are designed together with the poor or designed directly by the poor, within poor communities as they innovate on their own with new technologies (Heeks, 2009).

ICTs are found throughout the development process and in every area of development work. They support development organizations in improving their infor-

BOX 2

The potential of ICTs in the M&E cycle

Diagnosis. ICTs help bring new voices and broader participation into program diagnosis and enable a wider range of inputs at a reduced cost. They enable evaluators to better manage and pull possible trends out of large data sets.

Planning. ICTs can help achieve greater inclusion in planning processes. New technologies make it easier to compare and visualize data sets and to analyze data based on location so that resources can be better allocated. Data can also be aggregated more quickly and shared at various levels to improve participation in the planning process and support better decisions. New software tools can enhance the development and management of theories of change.

Implementation and monitoring. ICTs allow for the collection of real-time data on participant experiences, behaviors and attitudes, meaning that analysis can be conducted early on in the process and course corrections can be made to improve interventions and outcomes. Direct feedback from program participants is also possible through new ICTs, which can allow for greater transparency and accountability.

Evaluation. ICTs can increase the voice of vulnerable and underrepresented groups and broaden the types and volume of data that can be collected, combined, compared and analyzed. New technologies may be able to help overcome challenges and constraints such as sample bias and poor data quality, and they can improve the understanding of complex sets of behavior and data.

Reporting, sharing and learning. ICTs enable wider circulation of evaluative learning, interactive sharing and greater public engagement with evaluation findings.

mation management, public outreach, advocacy, influence and fundraising. ICTs are also used directly in programs, where they help people access information, markets, healthcare, financial services and education. They enable community members to connect with friends and family, and to augment their overall participation in the development process. Increasingly, ICTs are enabling improved feedback and participation from the populations that development agencies serve.

Those who were previously unheard in discussions about development are starting to use devices, software and platforms such as the Internet and mobile phones to enter into development debates and make themselves heard. ICTs have spurred innovative approaches to data collection, new combinations and comparisons of data and information, and faster data processing that facilitates better planning

and decision-making. The widespread availability of mobile devices means that information can be submitted from or collected in places that were difficult to reach in the past. In addition, people can share and communicate in new ways through these tools.

Earlier in this paper we discussed the broad context in which M&E is operating, the real-world and methodological challenges facing current M&E systems, and the difficulties of adapting traditional M&E approaches to a rapidly changing and increasingly complex international development scenario. While ICTs cannot single-handedly resolve all of the challenges listed above, there are some tools that can be used throughout the planning, monitoring and evaluation cycles to help overcome limitations in conventional M&E methods.

Despite a surge in activity in the area of ICT-enabled M&E, many evaluators still use traditional methods and approaches. While there are certainly cases where traditional data collection methods are most appropriate and ICTs create their own set of new challenges (as we will discuss in Chapter 5), an investment in the development, application and evaluation of innovative new M&E methods that include creative uses of ICTs could help organizations adapt their approaches throughout the entire program cycle, making them more flexible and adjusted to the complex environments in which development initiatives and M&E take place.

© InsightShare

3

Current trends and tools in ICT-enabled M&E

New software and technology devices and tools developed over the past few years have given rise to new approaches to M&E. In addition to voice calls, widespread use and ownership of basic mobile phones enables people to send and receive text messages in remote settings. Slightly more sophisticated mobile phones allow for installation of applications, such as surveys for mobile data gathering. In addition, these phones can take photos and record sound, and some have the capacity to track locations using global positioning systems (GPSs). Smart phones and tablets, a step up in sophistication, function like hand-held computers and can make mobile data collection easier and more intuitive. In locations with steady 3G, 4G or Internet network coverage, data can be constantly uploaded and stored in "the cloud". The cloud also allows applications to be updated more easily because software does not need to be installed directly onto a computer by IT professionals using a CD and complicated procedures.

Along with advances in mobile phone technology, an explosion of mapping tools, platforms, software and data visualization options offers greater possibility to combine data sets and support more informed decisions about resources and program implementation. Remote sensing and satellite imagery can provide bits of information that can be mapped for a better understanding of everything from whether water pumps are working to observing environmental degradation and large-scale migration due to conflict. "Dashboards" are being developed that make information available in almost real-time for program managers, donors and, in some cases, local government staff, frontline NGO staff and community members themselves. New tools are also helping evaluators manage M&E processes and results, including, for example, online theory of change software and "nano-surveys" that allow evaluators to collect survey questions from a random sample of Internet users.

Social media enters the mix by enabling broader discussion and engagement with information, including data that is shared more openly and evaluation information on good practice and lessons as well as failures. The ability of big data analysis to monitor social media communications provides a powerful evaluation tool. For example, after a radio program targeted at teenagers and discussing topics such as sexual harassment or the dangers of drug use, it is possible to monitor social media to identify increases in the number of references to these topics. Tracking

BOX 3

Defining monitoring and evaluation

While monitoring and evaluation have different purposes, the two are closely linked. Much of the information required for an evaluation will be generated through a monitoring system. Consequently much of the discussion of the potential applications of ICT for monitoring will also apply to evaluation, and vice versa. The following are definitions used for the purpose of this paper.

Monitoring. Monitoring is the ongoing collection and reporting on data during the process of project implementation. The data, which can be quantitative or qualitative, is intended to provide regular progress reports to managers and other stakeholders on program performance and to identify and address any problems during implementation.

Evaluation. There are two main kinds of evaluation:

- **formative evaluation** provides ongoing assessments of whether the program is on track to achieve its objectives and what kinds of corrections are needed, while
- **summative evaluation** assesses the extent to which the program has produced the desired outcomes and whether these can be attributed to the effects of the program.

This latter requires a process of interference through which an analytical model (such as a pre-test-post-test comparison group design) is used to exclude alternative explanations of the observed changes.

social media conversations makes it possible to analyze topics being discussed and how different groups discuss them. Some evaluators are using social media as a means to conduct focus group discussions or to track how program participants feel about a particular service or initiative.

There is an enormous range of possibilities for using ICTs in diagnosis, program planning, ongoing monitoring of activities, visualization of data for course correction and resource allocation, overcoming real-world and methodological constraints during evaluation, learning and sharing of evaluation results, and capacity building in the area of M&E. It is also possible to combine different ICT tools, mix traditional methods with new ICT-enabled approaches, and enhance the efficiency of traditional methods of data collection through ICTs. It all adds up to bringing about new notions of what is meant by "monitoring and evaluation", "research", and "data collection".

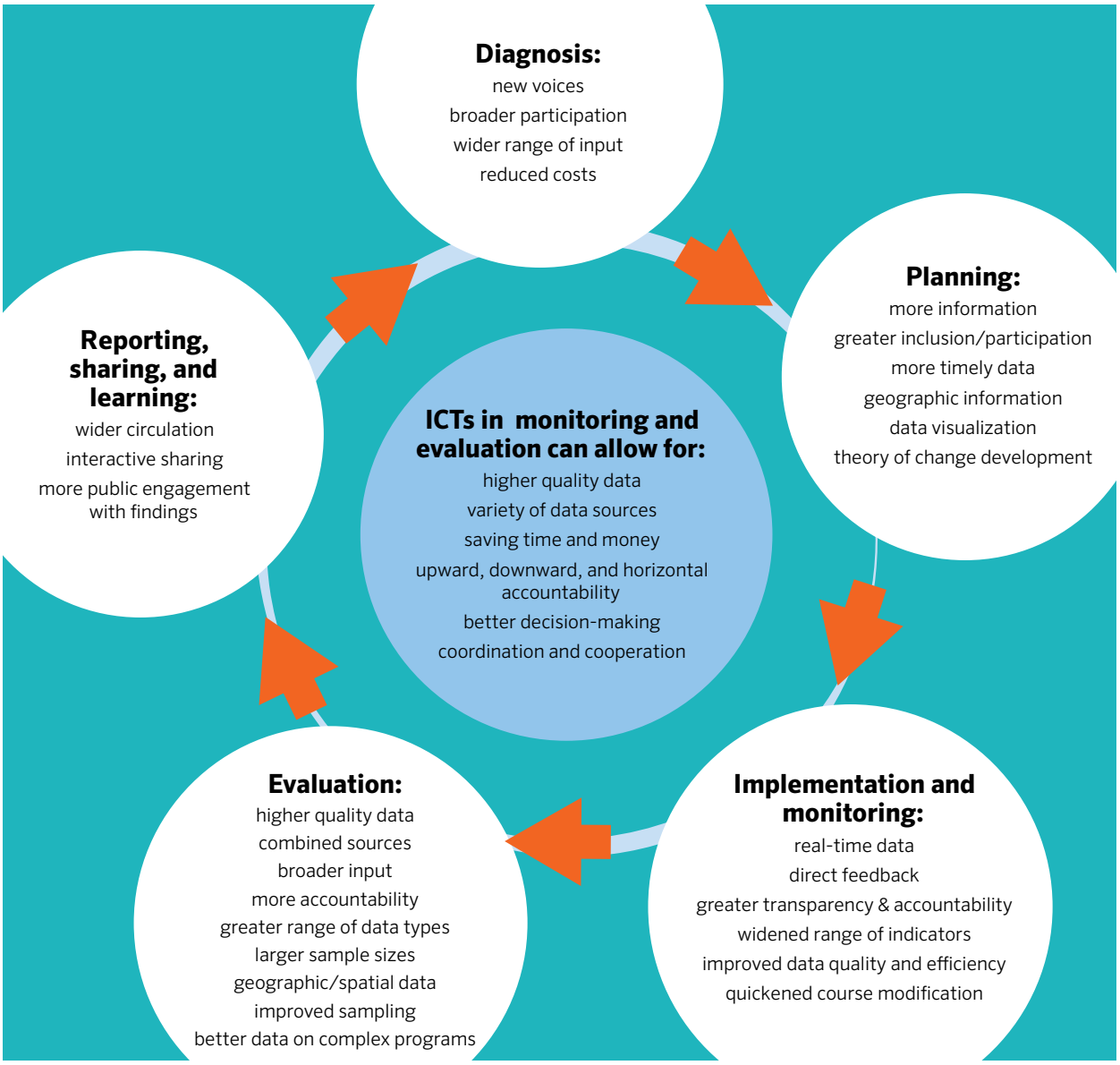
ICTs are being integrated into different evaluation methodologies and monitoring systems. They can support evaluators in collecting information that is of better quality in some cases, and there is great potential for new and varied uses of ICTs in evaluation. However, evidence is slim on how ICTs improve evaluation methods and processes, and there has been greater exploration of ICTs for monitoring than for evaluation. ICTs also present a number of new challenges, which we address in Chapter 5.

First, however, we highlight how ICT tools and approaches are being used for diagnosis, planning, monitoring, evaluation and learning. Though the tools are listed under particular stages, many can be used in multiple stages of the M&E process. They also can be combined and linked, and should be adapted according to the M&E needs and the context in which they are being used. We also recognize that M&E is an ongoing cycle rather than a linear process.

3.1 Diagnosis

ICTs are being used for diagnostic purposes at large scale in at least two areas: online or mobile phone-based consultations, and the capture and analysis of big data.

Figure 1. ICTs in monitoring and evaluation



Consultation

A growing criticism of many development initiatives is that interventions are designed and developed in offices in capital cities and do not reflect the priorities and needs of the people and communities they aim to support.

The expansion of mobile networks and SMS capabilities to even the most remote areas has spurred development organizations to consider how mobile devices can be used to engage the broader public in identifying and prioritizing issues that agencies should be addressing.

The United Nations, for example, initiated the "World We Want" campaign to encourage people around the world, especially those in developing countries, to input their ideas and priorities into the Post-2015 Agenda. In addition to more traditional face-to-face consultation with governments and civil

society organizations, the UN hosted online discussions on a wide range of themes and also conducted a campaign asking people to text in their priorities. The World We Want's website allows visitors to explore the responses through an interactive page, and data visualizations allow for easy comprehension of global priorities.[5] Because this approach can only reach those who have access to a mobile phone or the Internet, organizers of the campaign have also made an effort to combine online with offline consultation exercises and integrate them into the overall database.

BRAC frontline staff workers took advantage of regular meetings in communities to conduct a poll and send in community priorities by SMS.

The organization BRAC (formerly Bangladesh Rural Advancement Committee), for example, reached out to almost 12,000 village-level organizations in Bangladesh to ask community members what their priorities were. BRAC frontline staff workers took advantage of regular meetings in communities to conduct a poll and send in community priorities by SMS (May, 2013). The BRAC approach ensured that those without mobile phone access, literacy or funds to send an SMS were still included. UNICEF's U-Report, a similar consultation effort, worked with local partners in Uganda to engage over 100,000 young people as U-Reporters. UNICEF's Uganda office sent SMS polls to U-Reporters to gather their input, which was then used in determining program interventions or sent to government ministries to allocate resources to respond to health crises such as nodding disease (UNICEF, 2012). Conducting broad consultations using these new tools is helping bring new voices into the debate, and it is expected to help build greater ownership in the development process by a wider range of participants as well as lead to more relevant development efforts.

ICTs are also widening involvement in consultation and decision-making processes about M&E itself. Dispersed management teams and implementation teams are needed in any development monitoring or evaluation that extends beyond a local site. Tools – such as Skype and GoToMeeting, which offer inexpensive voice calls, conference calls and screen sharing, Google Drive and Dropbox, which allow large files to be shared, and Trello, which is a free task management application – help teams coordinate and broaden participation in the process of planning and managing M&E. These tools also support greater voice and engagement by people and organizations dispersed across sites, countries and regions so that M&E direction and decision-making is not centralized in one place.

Big data analysis

Big data is being used to conduct predictive modeling and to try to make sense of the behaviors of large populations or human systems and to forecast systemic shocks or changes more effectively at large scale. Big data can also identify idiosyncratic shocks and processes, such as when large volumes of creditworthiness data are examined. Big data is normally of higher volume, greater variety and quicker velocity, and it comes from a number of sources, including: sensors, social media sites, online photos and videos, online purchase records, mobile phone signals and call records, and other similar sources. Growing capacity to collect data related to people's actions and behaviors has prompted efforts to harness that data to predict and track behaviors and plan interventions more quickly than previously possible. In the past, often by the time a full-scale diagnosis of a problem or situation was conducted, it was too late for an effective response or the data were already outdated (Letouzé, 2014).

Large companies such as Google and Facebook have used large sets of marketing and user behavior data

[5] For examples of these data visualizations, see http://trends.worldwewant2015.org/discover/

to shape marketing efforts and earn revenue, but it is only recently that humanitarian and development agencies have begun to explore whether big data can be used to predict and track behaviors of those living below the poverty line. Global Pulse is one such organization working on research and experimentation aimed at finding connections between "data exhaust", i.e. the trails of data produced by those using the Web or mobile devices, and potential development interventions (Letouzé, 2013; UN Global Pulse, 2012). The Qatar Computing Research Institute is examining similar ways to track and filter relevant social media traffic for disaster response. Once developed, tools will likely be made available for development organizations to consider for their own efforts (Meier, 2013). Concerns have arisen about big data and privacy, given the increasing capacity to identify individual behaviors and geographic locations and trends, and this should be taken into consideration and carefully examined when working with big data.

A network called the Big Data & People Project (Data-Pop), created jointly by the Harvard Humanitarian Initiative (HHI), the Massachusetts Institute of Technology (MIT) Media Lab and the Overseas Development Institute (ODI), was launched in 2014. It brings together individual and institutional actors involved in the "big data revolution" to advance common principles and objectives. Data-Pop members are concerned with some of the claims that big data will lead to human and societal progress, especially for the poor. The network will more closely examine risks that come with the data revolution, such as the creation of a new digital divide and an overly technocratic and less humanistic approach to data collection, usage and decision-making (Data-Pop, 2014).

3.2 Planning

Data and input from constituents that aid the planning process are often gathered, accessed, analyzed, shared and discussed using ICT tools (as explained in Section 3 - Consultation). ICTs can also support planning. One example of the relationship between ICTs and improved planning is the use of a theory of change (TOC).

Theory of change

A well-developed TOC provides guidance on what information to collect and how it should be interpreted. A TOC articulates the processes and mechanisms through which program inputs are transformed into outputs and, in turn, transformed into outcomes and goals or impacts. The TOC also defines the economic, social, political, socio-cultural and environmental contextual factors that can affect program outcomes and helps define the kinds of data that must be collected on each input, output and outcome indicator. For M&E systems that collect data digitally or use ICTs, a TOC can help avoid the tendency to focus on the kinds of data that are easy to collect digitally.

ICTs can facilitate the collection of data required to populate and use a TOC in a number of ways. For example, integrated databases or management information systems (MISs) can allow data to be input into the TOC from multiple sources, such as the program monitoring system, agency records or records from other agencies. The integrated database may facilitate the collection of contextual data including, for example, economic indicators, migration patterns, school enrollment rates or disease rates. Maps and satellite data may also provide measures of the road networks and access of communities to, for example, markets, hospitals or towns. More specialized data can also be gathered and input using ICTs. The Rockefeller Foundation-funded Oxfam America Horn of Africa Risk Transfer for Adaptation (HARITA) weather-indexed crop insurance program in Ethiopia uses a TOC and collects satellite-generated rainfall data as part of its M&E efforts.

Software, such as Do View, enables the TOCs to be developed online in a participatory way through inputs from people in different physical locations. The software also summarizes figures so that every-

thing appears on one page, and it can also disaggregate models and present different components on different pages. The text in the boxes of the TOC can be changed into photographs or drawings when discussing the TOC with community groups that have low literacy levels, thereby making the tool useful in a variety of settings, contexts and with different populations.

Document organization and research

Reference management tools such as Mendeley and Evernote help with the collection of studies and documents. Information can be integrated from webpages, and users can search and sort through databases of academic documents. Researchers can use tags for easy categorization of documents across different folders and pull references into bibliographies. Some of these systems allow users to set up groups where they share research and evaluation documents, and notes. Other features allow team members to highlight information and files that might be of use to other team members. The Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) focuses on systematic reviews and research. Its EPPI Reviewer is an online software tool for research synthesis that analyzes and manages data in literature reviews. It is also used for all types of systematic reviews such as meta-analysis, framework synthesis and thematic synthesis. Through this software, researchers can search and screen bibliographic information, characterize studies using key words, conduct quality and relevance assessments and, finally, they can pull together numerical, empirical, thematic or conceptual syntheses.

3.3 Implementation and monitoring

Many examples exist of the use of ICTs for monitoring. These can include, for example, formal monitoring and data collection done by staff (such as qualitative or quantitative data collection using a tablet or mobile device); management information systems (MISs) with digital components that enable individuals to self-monitor (such as using a sensor that tracks movement or health, or texting reports into a central place); and built-in monitoring (when an individual's use of a digital device provides information that aids with monitoring behaviors). The following examples illustrate ways that ICTs are used in monitoring.

Incident monitoring

A clear understanding of the situation on the ground is often difficult to glean because staff and evaluators cannot be everywhere at the same time. In difficult political or crisis situations, staff cannot mobilize to more dangerous zones. Collecting reports or data directly from individuals or trusted sources living in a particular location through mobile devices means that the information can be transmitted almost immediately, also known as "in real-time".[6]

There has been growing awareness about the types of precaution that need to be taken to ensure that people are not put at risk by reporting sensitive or volatile incidents.

One of the first events using text messages and a digital map to collect and visualize incidents of election violence happened during the Kenyan elections in 2007 when the "Ushahidi" platform was developed to monitor election violence. This process is an example of "crowdsourcing", i.e. the practice of opening up a question or topic to a broad public for input or solution. Though gathering input from people is not a new method, crowdsourcing is made easier by new communication channels such as SMS and the Internet, which allow input from anyone who

[6] Global Pulse literature notes that real-time does not always mean "immediately". In international development, it can be understood as information produced and made available in a relatively short and relevant period of time, and information that is made available within a timeframe that allows action to be taken in response, i.e. creating a feedback loop.

has the motivation and tools to provide it. Two tools, FrontlineSMS and Ushahidi, are at the forefront of this trend and have been used together to crowdsource reports about a wide range of topics, from sexual harassment against individual women on the streets of Egypt, to incidents of violence against children in Benin, to projects that involve women in disaster risk reduction. [7]

Experience and use of ICTs for crowdsourcing has developed significantly over the past several years, and a number of lessons have been learned as to how it can be most effective. At the same time, there has been growing awareness about the types of precaution that need to be taken to ensure that people are not put at risk by reporting sensitive or volatile incidents. In addition to the more public type of crowdsourcing, these same tools are being used by frontline staff or other trusted intermediaries to capture reports when they are moving about in communities. FrontlineSMS and Ushahidi (and Ushahidi's simpler platform, Crowdmap) allow data that are collected to be exported to other software to conduct further analysis. These tools can allow for ongoing monitoring of topics, issues, sentiments and a range of other aspects, providing an additional set of data to accompany more formal monitoring processes or to raise areas of concern that require further attention or resources.

Social monitoring and feedback loops

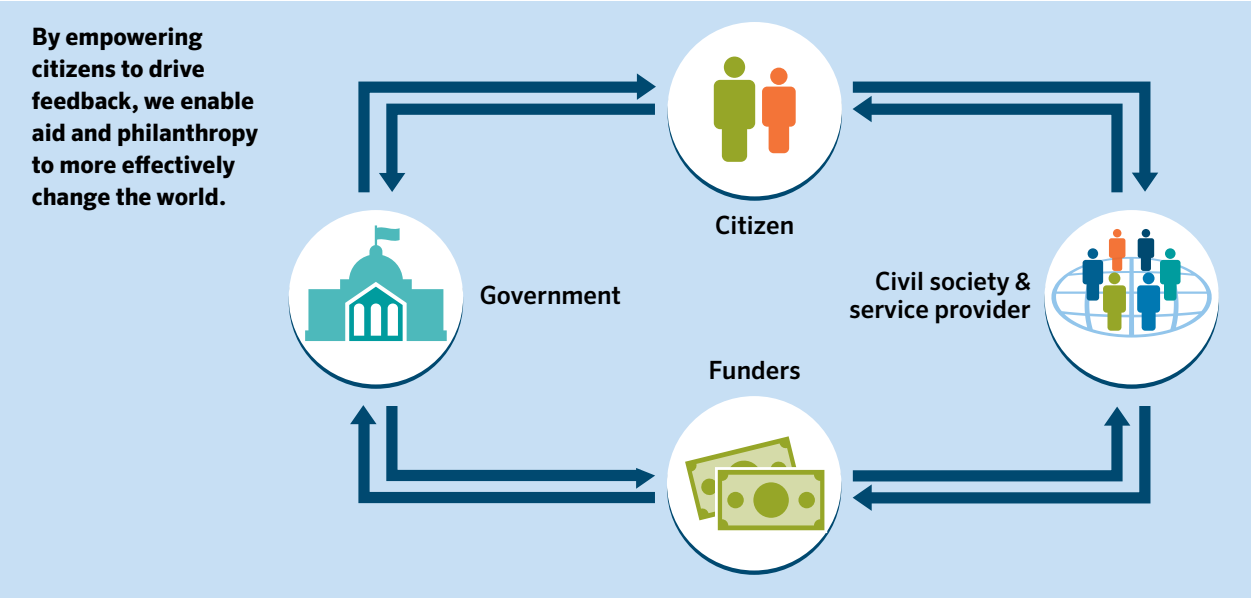
The narrowing communication gap between organizations and community members participating in development programs is giving rise to a number of new efforts to engage communities in providing input about the quality of implementation of these programs. The term "closing the feedback loop" refers to the exercise of collecting input from program participants and engaging them throughout the program cycle, ensuring that any information collected by or about them circles back to them and, at the same time, reaches those making decisions about programs and resource allocation. Experimentation with feedback loops through mobile phones (both voice and SMS), and crowdsourcing and social media platforms is happening in many organizations. Some believe that integrating mobile devices and SMS into existing paper-based feedback mechanisms could help to open up communication with a greater number of beneficiaries and reduce manipulation and intimidation from community members who threaten more disempowered groups and prevent them from providing their input into how humanitarian operations are conducted (Tonea, 2013).


© Linda Raftree

The infoasaid program aimed to improve communication with program participants in humanitarian settings in order to improve impact and services. A key component of the monitoring process was regular feedback from communities. A review of the program found that greater interaction with drought-affected communities, enabled by a community radio show and mobile phones, improved organizational understanding of communities' needs and priorities, and led to changes in how assistance was given. In two of the program sites, staff from humanitarian organizations reported that interac-

[7] For more information, see http://blog.harassmap.org/about/, http://vacbenin.ushahidi.com and https://womenandgirlsonthemap.crowdmap.com/

Figure 2. Feedback Labs is researching what makes for effective citizen feedback loops.



tive radio shows highlighted communities' needs and concerns from the communities' perspective, which often differed from opinions and assessments raised in meetings led by humanitarian agencies. In one site, the implementing organization became aware of poor sanitation conditions affecting specific schools, and refocused its efforts to address these issues. In other instances, listeners called into the radio program with medical questions, which alerted agencies and government authorities to potential disease outbreaks. An unexpected benefit was that isolated communities said they valued the ICT-based communication tools and feedback channels as a lifeline to the outside world because they typically felt excluded from outside sources of information and powerless to access rapid assistance. Introduction of mobile phones and solar chargers seemed to alleviate some of this isolation while also providing valuable information that enabled humanitarian agencies to adapt and improve their services more quickly (Chapelier and Shah, 2013).

Some feedback initiatives are linked with donor efforts to make financial flows to aid and development programs more transparent, and to involve beneficiaries as "customers and clients" who have a right to demand good service and impact from aid agencies and governments. They may also be used to detect fraud and to ensure that grants are targeted to the poor. For example, GPS can be used to verify whether persons who send reports are actually physically located in the place they say they are reporting from. Feedback Labs is a coalition of technology organizations and social enterprises that are researching, developing and testing effective ways to gather and use feedback.

Management information systems (MISs)

When data are collected and only sent "up the chain" to headquarters or, alternatively, when information remains local and is never consolidated or analyzed, it is difficult for communities, frontline staff, program managers, headquarters and donors to have a good handle on the ongoing activities and short-term results of an intervention. To resolve this issue, some large international organizations, including World Vision and Catholic Relief Services (CRS), have developed organization-wide M&E systems that capture information from a number of field offices. Information collected in an on-going way includes

agency indicators, geo-location of infrastructure and project implementation, activity plans and spending. A recent study found that some 55 percent of health programs incorporate some type of ICT for M&E, and 70 percent of health programs incorporate mobile devices (Bruce, 2013).

One advantage to programs and projects that use ICTs as part of their approach is that monitoring and tracking can be built into program implementation.

DevInfo, an MIS developed for UNICEF by Community Systems Foundation, enables collection and consolidation of global indicators on the wellbeing of children, placing statistics and country-level information at the fingertips of UNICEF staff. Similarly, DevResults, a web-based project management system, can be adapted to individual organizations and programs. It includes a dashboard to give managers and implementers a snapshot of program activities, financial metrics, results data visualizations, breakdown of project by sectors and status, and tracking of activities and indicators. Dimagi, a for-profit social enterprise, provides a choice of tools that can work together to manage health information systems at the level of frontline health workers and program managers. The tools support case management, data collection, data management, two-way messaging, surveys, logistics and supply chain management, stock tracking and delivery acknowledgement. Salesforce is a platform upon which to build customized M&E data collection, data storage and data management solutions. While ActivityInfo and Sigmah are open source MISs, their application to development evaluation is limited in terms of functionality and usability.

Increasingly, MISs are being designed to enable data collection from multiple types of devices or to allow paper, computer and mobile-based data collection (selected based on local context and conditions) to all feed into a central database. Information is being collected and sent up the chain as well as consolidated and made accessible to frontline staff, district- and national-level managers, and donors. In some cases, the information is made available to communities so that they can also benefit from it.

Selecting the right MIS can be a challenge. The selection can depend on a number of factors, such as whether to choose an open source system – many of which are not geared towards development work – or a tailored, proprietary system. Very few open-source MIS solutions aim at development or M&E, and those that do exist have limited functionality and usability. Because of this, organizations often opt to build a system that is customized to their project, program or organizational M&E needs. Any of these MIS options can require considerable investment because most development organizations do not have staff capacity to adapt and customize open source platforms, and proprietary systems often have high costs for training, installation and long-term service agreements.[8]

Built-in monitoring

One advantage to programs and projects that use ICTs as part of their approach is that monitoring and tracking can be built into program implementation. For example, if people access information or services via their mobile phones or the Internet, their online or mobile behaviors can be tracked and monitored to determine whether a program intervention is having its intended effect. The Akazi Kanoze youth employment program in Rwanda,[9] for example, has several components aimed at supporting youth workforce development and employment. Youth take online quizzes following training activities, and also participate in a mobile social network developed by an organization called Souktel, to access information about jobs. Through these access points, the implementing partner, Education

[8] Discussions with Jamie Lundine, Spatial Collective, May 2014.
[9] For more information on this program, see http://akazikanoze.edc.org

Development Center, can track whether students are learning, how often they seek job information, when they send in a job inquiry or resume, and how successful they are at landing a job (Education Development Center, 2013). Thinking about where these different touch points are with program participants can allow organizations to monitor programs in an ongoing way without spending too many additional resources on physically visiting and surveying participants.

> *Real-time or near real-time data which can allow for better decision-making about program implementation and enable implementers to conduct course correction earlier in the life of the program.*

In an effort to provide information to sex workers in Edmonton, Canada, about services they could access, the Center to End All Sexual Exploitation (CEASE) compiled a list of 25 anonymized phone numbers drawn from the adult services section of the "Backstage.com" website, and sent out pre-scripted text messages offering services and support. When the pilot was well-received, CEASE moved on to use a program called GSA Email Spider that automatically extracted phone numbers from Backstage.com, and added FrontlineSMS to manage the outgoing and incoming text messages, saving time and money and allowing for detailed analysis of the program in order to adapt it to make it more effective among the target population. For example, CEASE learned that 97 percent of replies to its outgoing text messages came in within the first hour after being received, and that those sent between 12:00 and 4:00 pm were the least likely to receive a reply. In this way, CEASE was able to adapt its scheduling and conduct more effective outreach (Gow *et al.*, 2013).

The "PartoPen", a multi-media pen that is used to fill out a paper "partograph," also allows for built-in monitoring. The WHO considers the partograph, when used properly, to be the most effective tool for monitoring women in labor and reducing labor complications in developing countries. The PartoPen interacts with the special ink used to fill in the partograph and, when tapped on different areas of the partograph, it provides maternal health training instructions, task reminders and audio feedback in real-time. The pen is also able to detect abnormal labor progression by analyzing data entered on the partograph form. In this case, it provides audio and text-based feedback to encourage birth-attendants to take appropriate action. Evaluators used the pen's capabilities to measure errors, corrections and all marks made on the partograph form to evaluate whether providing a tutorial on the PartoPen improved health workers' use of it (Underwood *et al.*, 2012).


© Adam Wills

Real-time data

One of the advantages of using ICTs for data collection in all of the above situations is the possibility of

collecting real-time or near real-time data which can allow for better decision-making about program implementation and enable implementers to conduct course correction earlier in the life of the program.

One program in Zimbabwe recorded a $10,000 savings by switching to tablets to survey a sample of 5,000 people, as compared with using a 25-page paper questionnaire.

3.4 Evaluation

As mentioned, ICTs have been used more for monitoring than for evaluation. Below we outline some tools and approaches that are being used more specifically for evaluation. However, many also touch on the other stages of the M&E process and are not exclusive to evaluation.

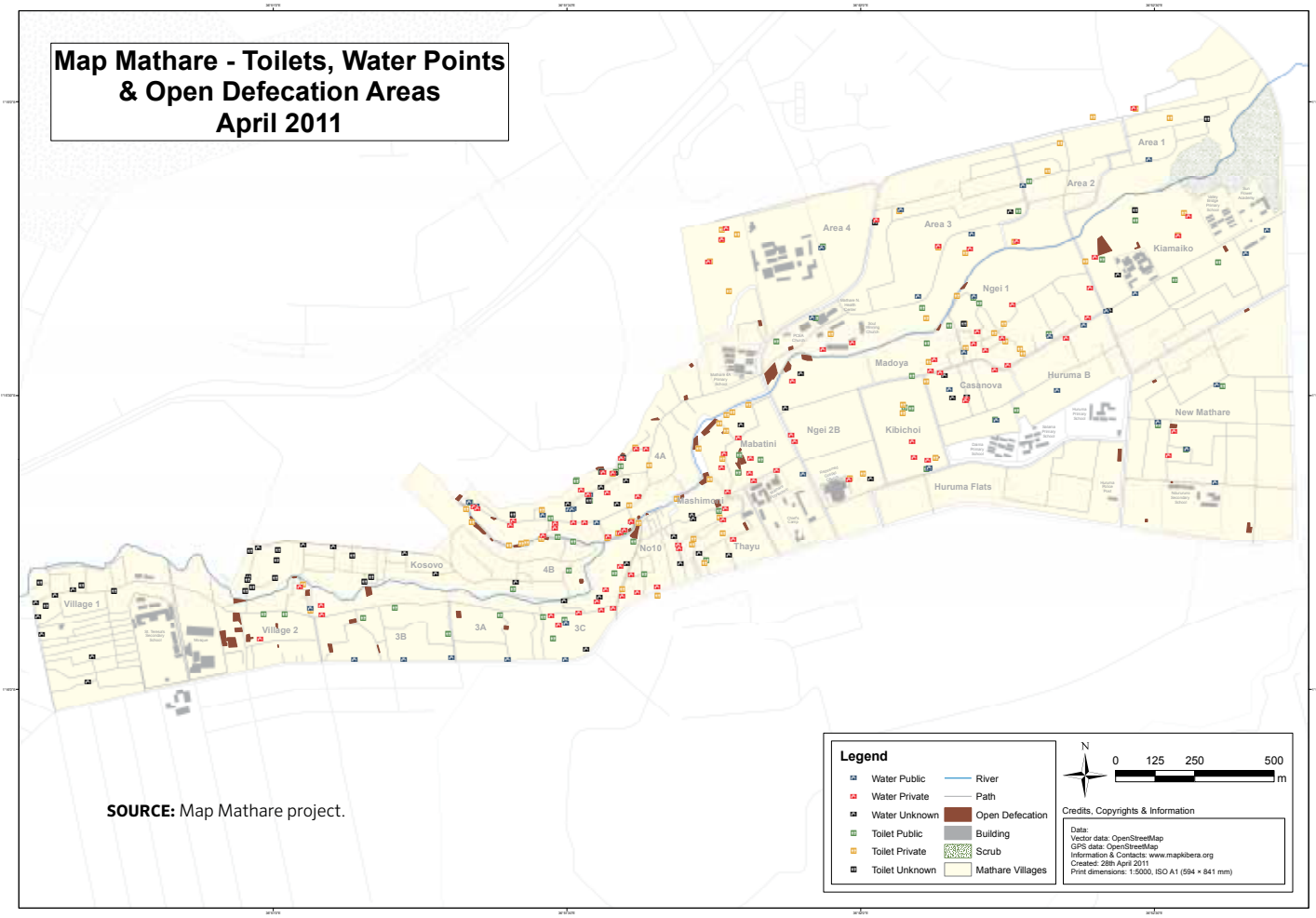
Data collection and surveys

Mobile data collection is perhaps the most well known use of ICTs in monitoring and evaluation. Traditional data collection using paper and pen can be time consuming, expensive and prone to error. Conducting surveys on a mobile device can save time and effort and, if the effort is well designed and staff well trained, mobile data collection can improve data quality. For example, a 2009 study found that none of the errors presented in 20.8 percent of paper questionnaires were found in the data set collected by a mobile device. Data entry, validation and cleaning for the mobile-based data collection from 120 participants showed a 93.26 percent reduction in time compared with that using paper and pen. Cost was also significantly reduced, and both data collectors and participants showed a preference for the mobile-based data collection process (Yu *et al*., 2009). One program in Zimbabwe recorded a $10,000 savings by switching to tablets to survey a sample of 5,000 people, as compared with using a 25-page paper questionnaire. Enumerators found it less cumbersome to carry around a tablet as compared with paper surveys. They were also able to collect data while standing or sitting in uncomfortable settings, and they reported being able to establish and maintain rapport with survey respondents more easily when using digital devices (Trigg, 2013). Though there are reported benefits to ICT-enabled data collection, there is relatively little solid research on its potential downside. Even less is known about potential issues such as the loss of data (data reduction) through transforming open-ended questions into multiple choice, or whether interviewers are less likely to follow-up multiple choice questions with probes than they might be with a paper and pencil survey.

Data collected through a number of different devices can feed into an organization's integrated database or management information system (MIS). The Regional Center for Learning on Evaluation Results (CLEAR), South Asia, has developed a guide that organizations can use to identify whether mobile data collection is a good choice for their M&E activities (Thakkar *et al*., nd). The World Wide Web Foundation has also conducted research on the use of mobile tools for data collection, with a focus on sub-Saharan Africa (Boyera and Alonso, 2012). Sambodhi, a monitoring and evaluation organization based in India, used mobile data collection tools to conduct a baseline study for the Rockefeller Foundation's Smart Power for Environmentally-sound Economic Development (SPEED) initiative and continues collecting data using mobile tools during the monitoring process.

In Vietnam, local community agents used the Open Data Kit mobile data collection tool to monitor forest management with the use of remote sensor monitoring of forest disturbances. It found that between 14 and 36 percent of the events identified by local community members were not detected by remote sensors and that, in some cases, remote sensors showed a delay of 1–2 years in capturing events. The role of mobile data collection by community members was highlighted as key to ongoing forest management and monitoring (Pratihast *et al*., 2012).

Figure 3. A baseline map in Mathare, Kenya, used to plan activities to improve sanitation



SOURCE: Map Mathare project.

Even when the data itself is not collected using a mobile device, cellphones have proven useful for coordinating and managing survey processes. The Busara Center for Behavioral Economics, a research laboratory in Nairobi, used FrontlineSMS to send bulk text messages to participants who had signed up to participate in research in order to remind them of their appointments. Instead of making 150–200 individual calls per day, a process that normally takes two field officers a full day to complete, one field officer needs only 30 minutes to send out the initial invitation to participants as well as a reminder closer to the date. The field officer can also customize messages to include different transportation reimbursement amounts, depending on the location of the respondent (Kuruvilla, 2013).

Collection of data through mobile phones given to survey participants is another potential approach. It was tested in 2011 as part of an experimental phone survey project conducted by the World Bank in southern Sudan. In this pilot, 1,000 households in 10 state capitals of southern Sudan were given mobile phones (Demombynes, 2011). Each month (starting in December 2010), Sudanese interviewers, based at a call center in Nairobi, phoned respondents to collect data on living conditions, access to services and citizen attitudes. Though the proliferation of mobile phones in developing countries has generated great interest in these types of surveys, a review of the initiative in 2013 noted that non-response was a substantial problem, mainly due to the erratic functioning of the mobile network. Response rates were

higher for those who owned mobile phones versus those who did not. Compensation provided to respondents in the form of airtime and type of phone was varied experimentally. Attrition was slightly higher for those who received higher compensation. The implementers of the experiment determined that mobiles phones can be a viable means of data collection, that calling people on their own phones is preferable to handing out phones, and that attention needs to be given to the potential for bias due to selective non-response (Demombynes, 2013).

... the use of "tagging" – meaning those who create the images or video determine key words that explain their main points – can help evaluators categorize the data ...

Mapping and geolocation

The widespread availability of global positioning systems (GPSs), including the inclusion of GPSs in some mid-range mobile phones, accompanied by increasing use of digital mapping platforms, such as Open Street Map, ArcGIS and MapBox, has led to increased incorporation of mapping into program planning, monitoring and evaluation cycles. Mapping is often used in several stages of the planning, monitoring and evaluation cycle because maps can be used to collect and visualize baseline data, to make decisions about where to allocate resources, to guide intervention strategies, and to visualize data collected during monitoring and evaluation to show results or gaps.

In some cases, large development organizations are collecting location data as part of the M&E process, though this can be costly if it requires staff training, additional equipment and more involved database management. Some mobile data collection applications, such as POIMapper and Open Data Kit, allow collection of a number of data points on a mobile device, including survey data, location information and photos. These can then be uploaded to a database where they can be visualized and analyzed.

The Map Mathare project used digital mapping as a tool for helping Mathare residents improve water and sanitation. Project organizers engaged community members in conducting a baseline survey and creating a digital map showing the incidence of public defecation. The community uses the map to make decisions about water and sanitation program activities, and the map will also help demonstrate results over time. Similarly, satellite maps can show large-scale changes in the environment, such as forest coverage, crop burning and even population movement, which can then be used to monitor activities, plan interventions or advocate at higher levels for policy changes.

The large international organization Pact uses maps in its work with community forestry groups in Cambodia in a program aimed at reducing carbon emissions through forest protection. In this case, satellite maps are used to generate awareness and discussion on how the forest has changed over time and to identify forest types. Community members also sketch out their own maps related to forest use, watersheds, timber resources, boundaries, conditions and conflict areas. These sketches are transferred into digital maps, which community members use to patrol the forest. They use GPS coordinates and photos to report illegal logging, endangered species sightings and land settlements. This information is sent to Pact by SMS, using FrontlineSMS. Involvement in data gathering on forest management allows tracking and mapping of poaching and illegal logging incidents, so that ongoing decisions around prevention activities and target areas can be made (Lamb, 2013).

Plan Cameroon, which works directly with children and youth in three communities, collects data on the location and quality of latrines, wells, trash dumps and water sources. The data are mapped, aggregated and analyzed by youth, who then print the maps and use them to discuss with their communities ways that improvements can be made. The youth


© Josh Nesbit

also use the maps and data to engage in dialogue with district government officials about budget allocation for water and sanitation. The baseline data will allow youth, with the support of Plan, to monitor and evaluate community and government follow-up and progress with regard to water and sanitation improvements (Plan, 2013). Mapping activities overall have become so popular in development work that the website Crowdglobe was established to collect and document examples and resources on the use of crowdsourcing and maps in development and humanitarian work.

Qualitative data collection

While there is a tendency to focus on quantitative data gathering when it comes to ICTs, the use of ICT tools to collect and analyze qualitative data is also being explored. For example, some mobile data collection applications allow surveyors, staff and communities to upload images, videos and recorded conversations. These can be collected and analyzed in ways that are similar to more traditional approaches. For example, the use of "tagging" – meaning those who create the images or video determine key words that explain their main points – can help evaluators categorize the data using software such as Sensemaker. Rating scales that assess social infrastructure by measuring the effectiveness and leadership styles of community associations and similar organizations can be enhanced by combining digital rating scores with audio recordings and photographs to compare ratings with what is observed in real life.

Organizations such as the United Nations Development Programme (UNDP) are integrating participatory video into their M&E activities in an effort to obtain more qualitative information as a complement to quantitative data. The organization Insight Share has adapted the Most Significant Change (MSC) methodology (Davies and Dart, 2005) by integrating participatory video.[10] In the original MSC method, mid-way through the M&E cycle, participants divide into story circles to share stories of significant change. Program participants select one story from each circle as the "most significant" and turn it into a testimony or act it out as a drama. Participants group key themes and analyze them during the evaluation, where the most significant stories are again identified and shared in an effort to maximize feedback and learning. Insight Share's methodology uses participatory video rather than oral stories, and community members create their own video stories throughout the process to explore and document key issues faced by the community (Insight Share, 2012).

ICTs are being used in randomized designs in a number of ways.

Strengthening randomized control trials

Randomized control trials (RCTs) are widely considered the most rigorous quantitative approach to impact evaluation. Their main benefit is that the random assignment of subjects (individuals, schools, communities) to the treatment and control groups

[10] MSC is a participatory monitoring and evaluation method based on narrative and storytelling.

can reduce problems of selection bias that seriously affect non-experimental designs.

ICTs are being used in randomized designs in a number of ways. For example, there is a large literature on how behavioral economics researchers randomly manipulate written communications from the government to the public in order to observe the impacts that wording variations have on public behavior.[11] When the government of the United Kingdom began incorporating text messaging into its toolbox of communication methods for collection of court-imposed tax fines, it conducted an RCT to test different approaches and compositions of the text messages. The test included six variations: not sending a text message, sending a standard message, sending a personalized message with the amount owed, sending a personalized message with the recipient's first name, and sending a personalized message with the name and amount owed. Responses were monitored to determine how personalization impacted response rate, time to payment and size of payment made (Cabinet Office Behavioural Insights Team, 2012). Similar RCTs were done to test the effects of different kinds of email messages on charitable giving (Cabinet Office Behavioural Insights Team, 2013a), and to determine how different messaging, pictures, and placement of messages and pictures would impact website visitors' inclination to become organ donors (Cabinet Office Behavioural Insights Team, 2013b).

Mobile devices are being incorporated into RCTs in order to address challenges with experimental social science research, including difficulties in measuring long-term impact, inability to observe treatment effects over time, inflexible research plans, high costs and close-ended survey questions. A project in Kenya run by the Policy Design and Evaluation Lab of University of California San Diego is introducing mobile phones to conduct surveys in order to augment an RCT that examines the impact of cash transfers to poor households. The project provides 360 respondents (180 treatment, 180 control) with mobile phones that they can keep at the end of the study (allowing researchers to follow up over time). Respondents are compensated via the mPesa mobile money service. The idea is that mobile phones can i) allow the researchers to test the impact of the intervention 12 to 15 months afterwards and also reduce attrition and cost, ii) allow high-frequency (daily, if needed) data collection, iii) make surveys more flexible because new questions can be easily added, iv) reduce cost because questions are sent by SMS and respondents are compensated via mobile money, and v) allow respondents to share thoughts to open-ended questions (Haushofer and Neihaus, 2013).

Social media monitoring tools and analysis of big data can help organizations measure how well this contributes to their thought leadership, alliance building and influence.

An RCT evaluation of a sports and HIV/AIDS prevention program in Cape Town is conducting a trial-within-a-trial activity to assess whether SMS reminders can enhance girls' knowledge, attitudes and self-reported sexual risk behaviors. A group of control schools receive standard HIV education. Within the intervention, half of the schools have been further randomized and their participants receive fortnightly supportive SMS messages. Trial participants self-complete questionnaires directly on mobile phones using the Open Data Kit survey software. Questions are in both English and isiXhosa, and can be either listened to or read (Ross, 2012).

Computer-assisted personal interviewing (CAPI), an approach used with RCTs to improve data quality and reduce cost of field data collection operations, calls for an in-person interviewer to use a computer

[11] For example, the Behavioral Insights Team (commonly known as the "Nudge Unit") reported that changing the wording in letters to people who owe back taxes can significantly increase overdue tax payment rates. See http://behaviouralinsights.co.uk for more information.

to administer a questionnaire to the respondent and capture the answers onto the computer. This interviewing technique became popular in the 1980s as computers became more widespread and less expensive. CAPI software is now available for tablets and smart phones as well as for laptops and desktop computers. However, a 2012 comparison of paper-based surveys and CAPI approaches noted that there was still limited empirical evidence on whether CAPI improved field operations and if it achieved better data than a well-designed, well-supervised pen-and-paper operation. CAPI requires basic data quality control and normally needs greater up-front investment than paper-and-pen, meaning that deciding to go this route requires careful assessment (Goldstein, 2012).

Nano-surveying, a technique developed by Real-time Interactive Worldwide Intelligence (RIWI), allows researchers to reach random samples of Internet users. The nano-survey repurposes the URL bar on the browser of any web-enabled device, creating an entirely new random contact point for data collection. When those searching the web type in a broken link, rather than getting an error message, they receive a mini quiz. RIWI's proprietary software and algorithms capture geo-location and eliminate fake responses to ensure geographical representation relative to Internet usage (including mobile, desktop and tablet) in any region.

3.5 Reporting, sharing and learning

ICTs and social media can support a number of the goals related to sharing and learning. They can also be used for training and capacity strengthening of evaluators and the field in general.

Thought leadership and influence

Some organizations have built an active presence on various social media sites and use evidence-based research and evaluation findings to talk about their work, linking to documentation and reports. Similar to how monitoring can be built into development programs, social media monitoring tools and analysis of big data can help organizations measure how well this contributes to their thought leadership, alliance building and influence. Orphio Technologies (formerly Media Badger) is a company that provides services to for-profit, public and nonprofit groups. The company analyzes publicly available data from websites, discussion boards and social media sites in order to provide insights into citizen and consumer views, including those on aid relief and assistance, and how well an organization is perceived by the public.

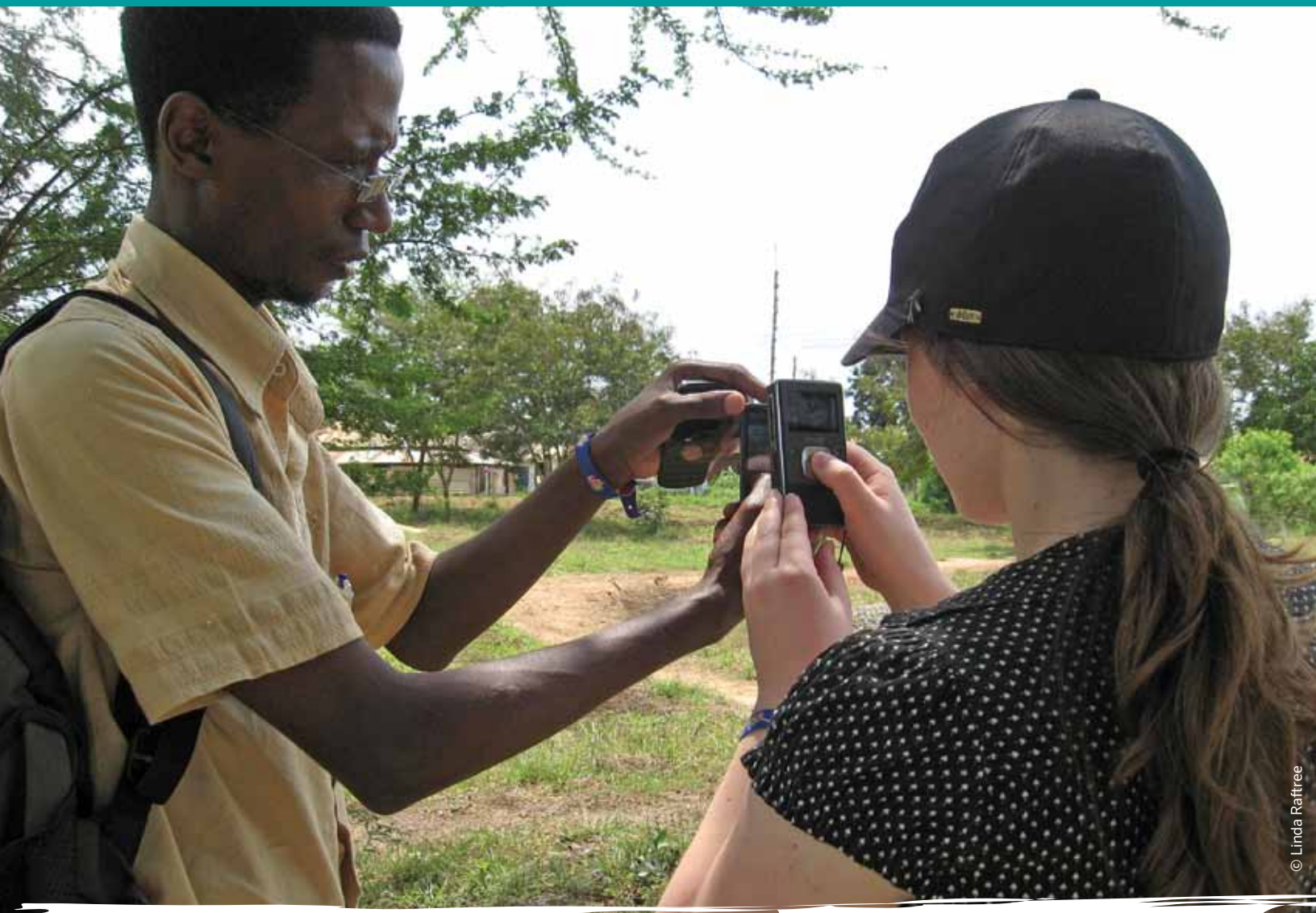
Making results accessible to practitioners and policy makers

Organizations and individual evaluation practitioners or firms can use blogs and online resource centers to make evaluation results more readily available to practitioners and policy makers. This helps bridge the gap between the academic and evaluation fields and those who can apply evaluation knowledge and learning to improve implementation and policy. Eldis, a website that is managed by the Institute for Development Studies, aims to share development policy, practice and research in a way that bridges the gap between practitioners and researchers. The site hosts free downloadable content from 7,500 development organizations and reaches 80,000 practitioners. Oxfam Great Britain's strategic advisor regularly blogs about studies, research and evaluation, providing analysis and insight to make these more accessible to practitioners and to share learning with colleagues at Oxfam and peers from other organizations and academic institutions.

Training for the evaluation community

A number of websites offer training and community forums as well as materials and resources specifically for the evaluation community. For example, contributors to BetterEvaluation - an international collaboration established to improve evaluation practice and theory by sharing information - curate and generate

information about evaluation methods, approaches and options, including guides and examples that cross sector, organizational and disciplinary boundaries. My M&E, an interactive web platform that provides knowledge on country-led M&E systems, aims to develop and strengthen a global learning and sharing community through blogs, discussion forums, documents, webinars and videos that can be readily accessed by M&E specialists around the world. TechChange, a young organization that offers a number of online ICT4D courses, recently conducted a 4-week certificate course on building skills and strategies to plan, collect, manage, analyze and visualize data using a variety of technology tools. The course features interactive education and learning, and guest presentations through Skype and Google Hangout with leading M&E practitioners, software developers and data scientists.



© Linda Raftree

4

Areas to explore

The above examples of ICT use in M&E have all been documented or are currently being tried out by evaluation practitioners. A number of potential uses also exist that have not yet been fully explored.

4.1 Random routes

Mobile devices could be used to improve the validity of methods such as randomly selecting households to be included in surveys. In the case of random routes, the smart phone could generate the instructions used to select subjects randomly and could also use GPS tracking to ensure that the correct household has been interviewed.

4.2 Reconstructing baseline data

ICTs could play a role in improving baselines. Although no precise figures exist, it has been estimated that perhaps as high as 75 percent of program evaluations do not begin until the program has been underway for some time, and frequently the evaluation is not commissioned until the program is nearing completion. Under these common evaluation scenarios, it is usually the case that no baseline data has been collected, making it difficult to apply pre-test-post-test evaluation designs. A number of tools and techniques have been developed to help reconstruct baseline data (Bamberger *et al.*, 2012a). The most common techniques include:

- reviewing available secondary data
- asking respondents to recall the situation at the time the project began (e.g. agricultural output and prices, time and cost of travel to school or work, expenditure of basic essentials)
- conducting key informant interviews
- holding focus groups
- using participatory group consultation methods such as participatory rural appraisal (PRA)[12] (Kumar, 2002) and most significant change (MSC).

While recall is useful and often the only available way to obtain information on the situation at the time of project launch, the disadvantage of this approach is that there are potential biases which are difficult to check. For example, do farmers under or over report agricultural production or do families under or over estimate other activities, such as the number of hours spent collecting water.

ICTs could be used to reconstruct baseline data. For example, when people enroll in a program (e.g. housing, microcredit, child nutrition, crop insurance), they could be given or loaned a phone and asked to take photos of, e.g. their fields, their children or the

[12] See Kumar, 2002, for an exhaustive list of PRA approaches.

kinds of food they have in their house. This could provide baseline data in cases where the organization has not been able or willing to collect conventional baseline data. Photos could also be taken of, e.g. the condition of roads or community centers. It should also be possible to teach women how to record the time they spend collecting water, preparing food, or caring for children, the sick and elderly. A stop-watch system could be used to record the start and end time of different activities or they could take photos. Travel could be recorded using GPS enabled phones. Applications could be developed that record visits to clinics, stores and other community facilities, or for people to record expenditures and photograph the products they buy or the bar codes of those products. Bar codes can collect information on the quantity and nutritional content of a product,[13] which is of use to researchers.

Existing studies on the importance of motivation and the impact of different kinds of incentive for information sharing could be reviewed to determine if and how best to encourage participants to provide information. It would also be important to review potential privacy issues, test the quality and veracity of the data provided, determine the type of training required for individuals to participate, and consider the varying levels of skill and literacy of participants in order to determine if these methods would be effective.

4.3 Improving sample design

Quantitative evaluations must define a way to ensure that the sample of subjects to be interviewed or observed is unbiased and representative of the population from which it is drawn. This normally requires at least a three-step strategy.

Step 1. Identify or use an existing sampling frame that includes all (or as high a proportion as possible) of the units in the population to be studied. This includes, e.g. individuals or households with certain characteristics, schools, community organizations or commercial enterprises.

Step 2. Use a sampling frame, to ensure that the sample selected is unbiased and represents the total population. Incomplete or biased samples can limit the validity of evaluation findings.

Step 3. Include a quality control procedure to ensure that interviews are conducted at the correct houses or locations and with the correct person.

There are a number of ways in which ICT can potentially strengthen sample selection.

A challenge in Step 1, which is common to all evaluations, not just those using electronic data, is that many existing sampling frames are incomplete. This may be because the sampling frame is out of date, it may intentionally select for certain characteristics, or it may unintentionally exclude important segments of the population of interest. A common and serious issue is that excluded sectors are often the poorest, most vulnerable or most difficult to reach groups. The challenge of Step 2 is ensuring that the appropriate individuals or sub-groups are selected from the sampling frame. Inadequate sampling frames can seriously limit the validity of evaluation findings.

There are a number of ways in which ICT can potentially strengthen sample selection. For example, through smartphones, interviewers could access real-time, updated registries that have the latest information and lists of, e.g. house-owners, credit union members or schools in the sample area. Online maps can assist interviewers with determining how many individuals, businesses, schools and other points can be identified that were not included in the sample. The proportion of new units identified is then used to estimate the proportion of under-

[13] Of course, this is only possible if food is packaged and has a barcode and available nutritional information.

representation of the original sample. Maps and GPS tracking can be used for quality control purposes to ensure interviewers are conducting interviews in the right locations. Where adequate sampling frames do not exist, samples can be selected using techniques such as random routes.[14]

4.4 Enhancing rating scales

Many evaluations use rating scales to rate quality of community infrastructure (e.g. roads, housing, drainage, water supplies). These ratings often must be subjective because the scales do not provide precise guidance on how to determine the correct rating, such as how to rate if the condition of a road or community center is "good", "satisfactory" or "poor").

ICTs could enhance rating scales by, e.g. using smart phones and tablets to provide photographs with examples of roads that are well maintained versus poorly maintained, or to create checklists with detailed guidance on how to use each scale. Combining these with geo-location information, for example, would make it possible to provide photos with examples of well and poorly maintained roads for particular locations. Ratings made by different data collectors could be checked online and, if there were inconsistencies, real-time feedback could be provided. Data collectors could send photos so that the study supervisor could check the ratings. For large-scale studies, it would also be possible to conduct inter-rated reliability assessments to identify potential biases or inconsistences. With these real-time approaches, corrections could be made immediately rather than at the end of the study, as is currently done.

[14] A random route sample is selected by giving interviewers a starting point in the community then giving them instructions such as: "Take the second street on the left and then select the third house on the right: then take the first street on the right and select the second house on the left." As it is usually difficult to check whether the interviewer is following these instructions (rather than skipping houses where no-one is at home or where the house is more difficult to reach, ICTs, including geo-location tools, could offer useful quality control.

Rating social infrastructure, which involves community organizational capacity and social networks, commonly calls for assessing the effectiveness and leadership styles of community associations and similar organizations. Rating scales can assess indicators such as leadership styles (e.g. participatory versus top-down), and compare male and female participation in meetings (not just attendance, but, e.g. who speaks, who is appointed to committees, whose views have more influence). Using ICTs for this kind of rating would enable combining digital scores with audio recordings and photos that could help qualify the rating.

4.5 Concept mapping

Concept mapping is a structured methodology for organizing the ideas of a group or organization (Bamberger *et al.*, 2012b). It can bring together diverse groups of stakeholders and help them

rapidly form a common framework that can be used for planning, evaluation or both. For evaluation, concept mapping is often used to organize meetings with experts or stakeholders. Each member of the group lists statements that reflect the dimensions or outcomes that he or she believes the program should achieve. Computer software is then used to integrate these statements into rating scales that can then be used to assess performance of the program on each dimension (Kane and Trochim, 2007). Concept mapping can also be used to document people's conceptualizations of different elements, their relative importance and their relationship to each other.[15]

Concept mapping is often used for:

- pre-test versus post-test impact evaluation designs (projects are rated on each dimension at the start and end, and may use control groups)
- post-test evaluation with experts rating changes that have taken place over the life of the project
- rigorous selection of case studies that will be explored in depth (experts rate each project on the scales and then cases are selected to include projects that ranked high and low on different dimensions)
- evaluation of complex programs where it is difficult to apply conventional impact evaluation designs
- evaluation programs that involve qualitative outcomes and impacts or processes of behavioral change that are difficult to measure
- large, widely dispersed programs with many different components and dimensions
- national and international policy interventions (e.g. UN or donor-funded policies to promote gender equality).

ICTs could strengthen concept mapping in a number of ways. For example, defining outcome indicators and development of the evaluation dimensions could be done online, software is available for statistical analysis, and experts could be linked via Internet or cellphone. Internet analysis could be personalized through video conferencing, and qualitative indicators could be included through audio and video recordings and photos. Expert ratings could be compared with ratings from other sources, such as human development and gender development indices.

[15] Proprietary software is available to implement concept mapping. More information can be found at http://betterevaluation.org/evaluation-options/concept_mapping.

4.6 Evaluating complex development programs

Evaluation of complex programs[16] is a rapidly emerging topic in development evaluation (Forss *et al*, 2011; Funnell and Rogers, 2011; Furubo *et al*, 2013). Complex programs frequently involve multiple funding and implementing agencies, multiple components, multiple outcomes and multiple causal paths.[17] A key difference between complex programs and complicated programs is that for the former, causal paths are often non-linear and outcomes have a high degree of unpredictability (see Box 1 for a discussion of the differences between a simple project, complicated programs and complex interventions). There is growing recognition of the limitations of conventional evaluation methods for the evaluation of complex programs but, as yet, no standard approaches have been developed that are equivalent to RCTs or quasi-experimental designs.

Many evaluation texts refer to the need to adapt systems theories, but only limited progress has been made in operationalizing these theories. One participatory application of systems dynamics is a systems-based evaluation approach (Groves, 2013) that has been applied in the evaluation of the Zambia Antiretroviral Treatment (ART) strategy for addressing HIV/AIDS. The method articulates the interrelation-

[16] For an overview of complex program evaluation see RealWorld Evaluation (2012) Chapter 16.

[17] The Ontario Smoking Cessation Program (Smoke-free Ontario strategy) is a good example of a complex program. See Schwartz and Garcia (2011) in Forss *et al*., 2011 Mara and Schwartz (op. cit) for a detailed description of the characteristics of this program.

ship of real-life factors including perspectives and boundaries and how they affect implementation of the strategy. The approach is structured around three system concepts: interrelationships, perspectives and boundaries. The two main forms of data collection are:

- transformative: facilitated in-depth participatory dialogue to understand and improve the system for and by stakeholders, and
- representative: observation and investigation to produce a description of the system.

The two forms of data are integrated to facilitate a participatory four-stage process moving from i) aligning pre-understandings, through ii) developing a shared comprehension of the dynamic situation and iii) providing space for arguments on the validity of data and the re-interpretation of inter-relationships, perspectives and boundaries to iv) a move to action. A model is developed to describe the implementation process and then used for engaging stakeholders in dialog and subsequent data analysis throughout the program cycle.

The inclusion of different types of ICT-enabled feedback loops can allow for inclusion of real-time input on programs, their relevance and their short-term effectiveness.

Evaluations also face the challenge that complex programs often involve large numbers of different components, each of which uses different implementation strategies and generates different sources of data. Implementation strategies frequently change and there are often problems of communication and coordination among agencies and programs.

For all of these reasons, ICTs may offer some important contributions. For example, many complex programs involve processes of behavioral change that are difficult to capture with conventional methods. Photography, video, voice recording and behavior monitoring using mobile phones may be able to help. Complex programs generate many different kinds of data, which often makes it difficult to integrate or to access many of these sources of information. ICTs may be able to make data more accessible through integrated databases and hand-held devices. For example, Ministry of Education officials in parts of India could access information on the local education offices they are visiting (e.g. numbers of staff, functions, budgets, student enrolment) making visits more productive.

Big data is also a potentially valuable resource for evaluating complex programs, and new technologies may be able to help channel this data for assessing complex programs. For example, monitoring changes in text messages, social network communications on issues relating to gender relations or drug use could help evaluators understand the broader picture or identify areas for further research.[18] Big data can also organize data on climate and other environmental conditions, cropping patterns, migration patterns and other big picture data that can provide context on factors influencing complex programs. In short, complex program evaluation involves many elements of systems analysis which new technologies may be able to address. The inclusion of different types of ICT-enabled feedback loops can allow for inclusion of real-time input on programs, their relevance and their short-term effectiveness. In turn, this information can bring a greater understanding of how different elements of a complex program are working and where the intervention or approach may need to be adapted for improved performance.

4.7 Quantitative case study methods

The increasing importance of complex programs and the challenges of their evaluation have increased

[18] It should be noted that the ethics around accessing and using this type of personal data need more thought.

interest in new and broader applications of case studies. Case studies have traditionally been viewed as a qualitative approach that uses a relatively small number of cases to illustrate the different typologies generated in quantitative survey analysis. However, over the past few years, new approaches have been developed that permit quantitative analysis of studies that typically include up to 30–50 cases. Many of these approaches, such as configurational case analysis, consider each case (e.g. individual, household, school, community) as a unique microsystem with its own internal dynamic. Each case has different components, such as different family members, different parts of the agricultural production system, and different elements of a school. Each member of each element affects the program-related outcomes in its own distinct way. In addition, the interactions among all members also influence the outcomes. Cases also interact with other elements of the system being studied (e.g. the district education or health system, the microcredit system, the local farming system), as well as interacting with other systems (e.g. economic, political, socio-cultural).

Configurational case study analysis creates a matrix describing the characteristics of each case and records the degree to which each program outcome has been achieved for each case (see Table 1 for a simple example of a matrix). The identified configurations of characteristics are always or frequently associated with the presence or absence of outcomes. The analysis can be done manually for small samples with simple descriptions of cases, or statistically for larger numbers of more complex cases. The analysis identifies the conditions that are necessary and sufficient for outcomes to be achieved (Byrne, 2009).

Looking ahead, it is likely that case-based methods will have increasing importance for program evaluation, particularly for complex program evaluations. ICTs may be able to assist in implementation of these approaches. For example, it should be possible to build software into a device, so that field workers (or researchers) can input data into the matrix, which can then be analyzed online, providing feedback as to which families/groups, etc., are likely to be most/least successful in achieving certain outcomes. Real-time analysis would permit the matrix to be modified and could indicate extra data that is required on the spot, rather than having to wait weeks or months for data to be analyzed.

5

New challenges and risks when integrating ICTs in M&E

As shown in this paper, there is potential for ICTs in a number of areas within planning, monitoring, evaluation and learning processes, and ICTs are being experimented with in numerous exciting ways. However, a number of challenges and risks need to be kept in mind when it comes to planning for and implementing ICT-enabled M&E, in terms of both evaluative approaches and capacities.

5.1 Selectivity bias

The use of ICTs brings special challenges of selectivity bias. On the positive side, mobile devices or the Internet are useful for reaching and interviewing groups that are excluded from surveys due to the cost or time required to reach them, due to security risks for interviewers to enter certain geographic areas, or because some groups are reluctant or forbidden to speak to interviewers (for example gang members, people who are HIV positive, women who may not be able to meet with outsiders). On the other hand, ICT technologies can introduce other sources of sample bias because ICT sample selection methodologies are dependent on people's access to mobile devices or other ICTs.

Accessing ICTs requires resources, such as funds to spend on airtime, Internet access or electricity. In addition to simple access to a device, the poor and marginalized are often excluded from ownership and use for financial or social reasons. Cultural norms, language abilities and literacy levels also affect a person's access to and use of ICTs. A program in Ghana, for example, noted that 80 percent of women participating in an SMS-based health education program required someone else to read and translate messages for them. In addition, some 20 percent of the world's population lives without reliable electricity, meaning that keeping phones charged is a challenge and affects how consistently reachable they are (Farmer and Boots, 2013).

In addition to basic literacy, cell-phone literacy can be an issue for those who do not know how to use new technology. For example, some programs allow people to access information or participate in different languages or use voice menus, but they still need to understand messages such as "Press #1 for Kiswahili." Lack of access to and control over digital devices may mean that some groups, such as women, girls, children, the disabled, the elderly and

those with low literacy, are less able to participate in M&E processes that rely on them. Concerns about possible negative outcomes or breaches of privacy and trust can mean that some individuals participate less than others. Physical security can also make ownership and use of certain devices problematic, especially for vulnerable groups or those who live in conflictive or high crime areas. Somewhat different, although related, selection bias issues may exist with passive big data collection, as not all members of the sampled populations use the social network or other sampled sites or media.

Most nonprofits do not have the funds or the capacity to adapt open source software to their particular needs.

Crowdsourcing has important potential issues of selectivity and exclusion. Some organizations use crowdsourcing and SMS to identify program priorities and to obtain feedback on the strengths and weaknesses of ongoing programs and policies. If these channels for input are not combined with other ways to gather feedback, issues of exclusion become even more serious. In addition, when working with tools or approaches that rely on crowdsourcing, it is difficult to know how representative sampling is. The more information requested from participants (e.g. age, sex, location), the less likely participants are to continue participating in mobile or online surveys. All of the above can lead to wealth bias, urban bias, gender bias, age bias and other biases in terms of selection and representation of samples and input (Raftree, 2013a). There may also be sample bias if the right people are not sampled due to the ICT tool selected for the data collection, especially if the methodology requires self-reporting.

5.2 Technology- and tool-driven M&E processes

Because ICTs make it easier to collect certain types of data in a less expensive way, there is potential that the technology, rather than the M&E design or data needs, will drive the kinds of data that are collected. When tools drive the process, it may lack a theoretical framework to help identify what to measure and implementers may not think carefully about whether they are asking the right questions in the right way or testing key assumptions. The difficulties of using a small keyboard, for example, may mean that multiple choice and numerical data are preferable when gathering data using mobile phones (Raftree, 2013b). Enumerators using tablets for collecting data in Zimbabwe tended to collect less detail when open-ended questions were incorporated into questionnaires designed for digital data collection, compared with paper-based ones (Trigg, 2013). Big data relies on automation, meaning that quantitative data is more likely to be collected. Some tools and methods have been created to help collect large-scale qualitative data, such as tagging videos, stories and narratives, and using key words to sort through and organize responses. However, at times, these have proven too complex or time consuming. In addition, questions remain about the validity of results and whether gathering quantitative data through stories can be considered M&E (The Rockefeller Foundation, 2012).

Another problem with M&E processes that are driven by ICTs rather than by good M&E design occurs when systems are designed by and tied to particular software providers. It may prove costly to make changes and adjustments to proprietary software, and the data collection may end up being locked into a structure that does not readily adjust to changing realities. At the same time, as mentioned earlier in this paper, most nonprofits do not have the funds or the capacity to adapt open source software to their particular needs.

A major challenge, particularly for agencies that have limited experience with data management, is that ICTs make it very easy to collect information from many different sources. There is a risk that the data will be under-utilized because they are not linked to each other. A similar problem exists with non-ICT data collection but the risk is multiplied when ICT is incorporated due to greater ease with which data can be collected. As with any kind of evaluation data collection, before data collection begins, it is important to take the initial step of developing a plan that identifies the kinds of data that are required, the specific indicators, how data will be collected and analyzed, and how they will be used. The plan would also need to discuss how different sources of data can be linked so that parallel data streams are not created.

5.3 Overreliance on digital data

The ease and relative low cost of collecting data using ICTs may mean that conventional (time-consuming and expensive) quality control measures are not used. Agencies may become obsessed with the cost savings of using digital devices for data collection and emphasize data over personal contact and direct observation. Over-reliance on digital data collection can mean that evaluators miss the opportunity to spend time on the ground, walking around and conducting informational conversations with staff and community members. They may only focus on what is captured on the smart-phone, running the risk of ignoring contextual factors such as indicators of wealth, family relations, community dynamics, leadership styles in meetings, and participation of women. This can reduce the data to meaningless numbers, with little context.

Organizations may not have the skills and expertise needed to use ICTs to support their M&E goals.

When quantitative data is overemphasized, staff may learn to "game the system" by checking boxes to report "satisfactory progress" in order to avoid follow-up from headquarters (Raftree, 2013a). One organization that gives cash grants without any strings attached (unconditional cash transfers) uses feedback from smart phones to check that a rural family really does have a thatch roof (the indicator used to select poor families), and then checks GPS location to ensure the persons live where they claim).[19] It may not be too long before people learn how to fool the system in order to receive cash transfers for which they are not eligible.

5.4 Low institutional capacity and resistance to change

Organizations may not have the skills and expertise needed to use ICTs to support their M&E goals. For example, they may lack experience in how to select the right combination of ICT tools for their particular M&E approach. Estimating the long-term benefits versus the total cost of adopting a new approach requiring additional hardware, software and training can also be difficult, and delays may result when transitioning to a new system. A 2008 study found that both domestic and international nonprofit or-

[19] This is the system used by "Give Directly" http://www.givedirectly.org/ but many other organizations use similar systems


© Linda Raftree

ganizations cited issues such as technology, training, budgets, program fit and unstated, yet implicit, organizational culture barriers to using ICTs (in this case, personal digital assistants (PDAs) for data collection (Banga *et al.*, 2009). At the institutional level, the use of new ICTs requires new skills and behavior changes as well as an understanding of applying data and data management, privacy, protection and security to new settings, new devices and new approaches to data collection. Organizations may struggle to build capacities to take advantage of the benefits of ICTs in the best way possible. Staff may also be afraid to try new approaches due to fear of failure. At the same time, donors may push for innovation in the use of ICTs without a clear understanding of the challenges, ongoing costs and need for capacity strengthening. This can lead to prescriptive solutions rather than a focus on building local institutional capacity to make decisions about integrating ICTs where they will be most appropriate and relevant to the local context.

As with any process of innovation and change, a number of factors impact the successful integration of ICTs into the M&E process (Raftree 2012a, 2012b, 2012c, 2013c). Sufficient time for training needs to be built into the M&E plan, to ensure that all members of the research or M&E team are well-versed and have practiced using a new device or approach (Trigg, 2013). In addition, the goal should be for local staff to understand and use new technologies for monitoring without requiring constant support, supervision and intervention from international staff and consultants (Walker Hudson, 2013).

5.5 Privacy and protection

Many organizations and technology providers are unaware of the ethical implications of collecting data via new tools and channels, or the nature of the privacy and protection risks that come along with new technologies. This results in security, privacy and confidentiality not being adequately addressed. Many development organizations are unclear about the ethical standards for research versus information or data that is offered up by constituents or "beneficiaries" (e.g. information provided by people participating in crowdsourcing or SMS-based surveys) versus monitoring and evaluation information. It is unclear what the rules and standards are for information collected by private companies, with whom this information may be shared, and what privacy laws mean for ICT-enabled M&E and other types of data collection. This can pose an ethical challenge to evaluators who wish to use new ICTs for data collection and do not know where to find guidelines or orientation to help them ensure ethical and privacy standards are met. It can also lead to unintended negative consequences.

Use of ICTs for data collection and M&E can put staff, evaluators and program participants at risk. A researcher working in Zimbabwe noted that the use of electronic devices attracted attention and aroused more suspicion than the use of paper forms, and that enumerators were more likely to be accused of being journalists when they used electronic recording devices (Trigg, 2013). Similarly, it was noted that in highly militarized and politically charged environ-

ments, the use of technology for data collection can undermine the operation itself because smartphones attract attention and are often associated with intelligence gathering rather than needs mapping (Tonea, 2013). A careful, in-depth informed consent process is critical, even when working in less sensitive environments. When working with participatory video for M&E, for example, it is important to ensure that participants fully understand the implications of sharing their voices and opinions via video, that they decide on content and the shape of the final video product, and that they fully understand who may have access to the video now and in the future (Muniz, 2013).

6

A checklist for thinking through ICTS in M&E

The pros and cons of integrating ICTs into M&E processes are quite balanced and, with careful thought and planning, ICTs can be used with success. Here, we provide a checklist to help evaluators begin thinking about how they might include ICTs in their M&E design.

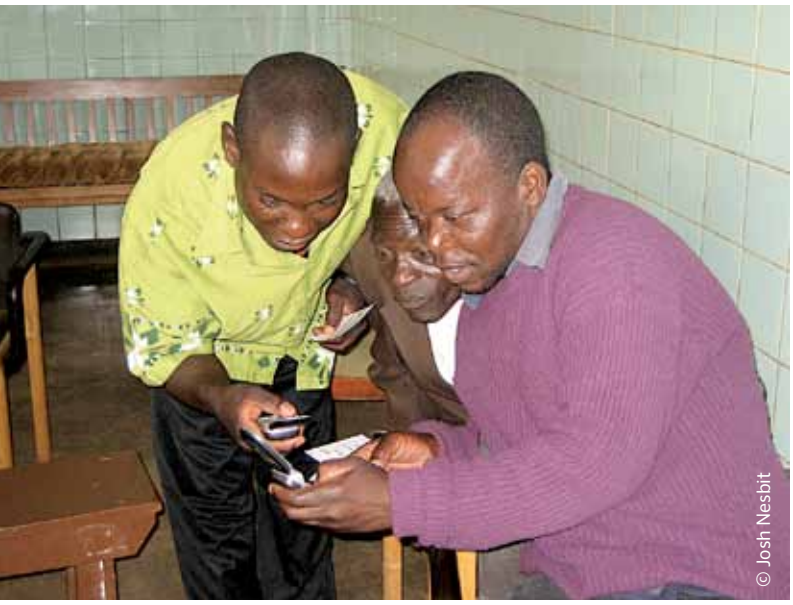
1. Develop a quality M&E plan

Adding new technologies to poorly designed monitoring and evaluation plans will not be of much benefit. However, a well-crafted M&E plan might benefit from the addition of ICTs, especially if the ICTs and the M&E plan are integrated from the very start of the initiative. Having a clearly articulated theory of change can provide guidance on what information to collect (with or without ICTs) and how it should be interpreted. Clarity as to the level and type of M&E that will be conducted can help identify the right kinds of ICTs to consider: Are you monitoring a project, a program or a wider initiative? Are you monitoring service delivery or influence? Are you more concerned with quantitative data or qualitative data or a combination of the two? What evaluation methodology will you be using?

2. Ensure design validity[20]

When designing evaluative processes with ICTs, it is important to be aware of the categories used to judge strengths and weaknesses of evaluation design, analysis and interpretation. The four categories that judge validity include: i) internal design validity: reasons why conclusions about the cause and effect relationship between two variables may not be correct, ii) statistical conclusion validity: reasons why conclusions about the statistical association between inputs and outcomes/impacts may not be correct, iii) construct validity: reasons why the constructs used to measure inputs, processes, outcomes and impacts may not be appropriate (e.g. income may not be a good construct to assess household wealth), and iv) external validity: reasons why assumptions about how generalizable findings from a pilot project are to different contexts may not be correct.

[20] Bamberger *et al.* (2012) RealWorld Evaluation has a chapter on assessing the validity of evaluations and several appendices with checklists that can be used to assess the validity of a particular evaluation.


© Josh Nesbit

3. Determine whether and how new ICTs can add value to an M&E plan

Although it is common to start with the technology and ask what can be done with it, one should start with the M&E plan and ask where new ICTs can add value or help improve the design, who the different stakeholders are in the M&E process, and what type of ICT would be most useful. In addition, the fact that more data can be collected more quickly through the use of ICTs does not mean that all the data will be of use. Organizational capacity is needed to analyze data as is the will to use the data for adapting and modifying program approaches according to what is learned.

4. Select or assemble the right combination of ICT and M&E tools

No single ICT or M&E tool is likely to offer everything that an organization or evaluator is looking for. In addition, most M&E specialists do not know what questions to ask when trying to find the right ICT tools. Hidden costs, technical support and training needs must be worked out in order to determine whether the return on investment makes integration of ICTs worthwhile. If the M&E goals are clear, a number of applications, tools and devices can be used to collect different types of information that feed into the overall process. Creating brand new, bespoke tools and applications for a specific M&E process may not be advisable, given the range of available applications. It might be best to consider systems that gather data in common formats that are easily shared, have available support for ongoing maintenance and conform to open data standards. However open source also requires resources and the capacity to adapt it to an organization's needs.

5. Adapt and test the process with different audiences and stakeholders

The right combination of ICT and traditional tools will depend on who will collect the data and who will use the data. Understanding the context, connectivity and capacity of these different audiences is critical during development of the ICT-enabled M&E process. Testing tools and data collection early in the process with a group of users can help identify areas where adjustments to tools, applications and processes are needed before conducting large-scale data collection or roll-out. As part of this process, keeping in mind the questions: "M&E for whom?" and "M&E for what?" can help ensure that the data collected meets the needs of the various stakeholders, whether they are donors, community members, program managers, government, or a combination of these and others.

6. Be aware of differing levels of access and inclusion

One of the premises of including ICTs in the M&E process is that they can help expand access and promote greater inclusion for the most marginalized members of the community. However, challenges, such as irregular access to electricity, poor connectivity, the cost of devices and providing content in local languages, mean that inclusion is not a given. As with any kind of community process, a key consideration is how to manage power dynamics and ensure that everyone has an equal chance to participate and provide input. Combinations of M&E tools and channels should consider local context, levels of access and network coverage of those collecting

and using the data. Context assessments should also include an understanding of how and whether communities use different kinds of ICTs (e.g. is voice or SMS preferable if relying on mobile phones), the languages and scripts that are available on different devices, which sources of information different groups within the community trust, and cultural concerns such as attitudes about women and girls' use of technology (Walker Hudson, 2013).

Smart phones may allow already powerful members of society to increase their power.

7. Understand motivation to participate in M&E activities

Motivation and incentive can impact the timeliness and quality of data collected during monitoring and evaluation. When data are being collected from community members, it is important to develop ways to feed the data back to them so that the data collection process will not be only "extractive". The same is true for frontline staff who are expected to provide M&E data. Opening ownership and sharing M&E data so that different stakeholders can use it for decision-making can help improve responsiveness, but privacy and risk need to be analyzed first. When broadening the M&E process and using approaches such as SMS reporting and voluntary "crowdsourced" input, a good understanding of motivation is critical. As found in one 2012 study, participant motivation, not technology, is the biggest constraint to effective crowdsourcing (Findley *et al*., 2012).

8. Ensure privacy and protection

The range of new tools available for data collection is wide, but those collecting data may not be aware of new privacy and security risks that come with them. Little documentation is currently available for those wishing to improve their understanding of the potential risks with ICT-enabled approaches to M&E. Yet, maintaining secure databases and taking care that digital data is protected is extremely important, especially when working with vulnerable populations, or in situations where corruption may be present or where conflict could be exacerbated by the M&E process. Having very clear and updated informed consent processes in place is critical.

9. Try to identify potential unintended consequences

In the push to promote innovation and ICTs, advocates promoting the benefits of emerging technologies tend to ignore the potential for unintended consequences that can result from introduction or use of ICTs. Domestic violence has been shown to increase in some cases when women are given mobile phones and men fear that they may use them to develop relationships with other men (Kutoma, 2010). Smart phones may allow already powerful members of society to increase their power (Stahl *et al*., 2010). Involving local staff and communities in an assessment of potential risks and keeping a close eye on what is happening outside the actual M&E work is critical for identifying potential unintended consequences and addressing them quickly if they happen.

10. Build local capacity

Often ICT-enabled initiatives focus on top-down, externally created "solutions" rather than building on local systems and processes, or working with local partners. Increasing the participation and improving the capacity of local evaluators and local partners is needed for high quality, sustainable M&E. In addition to finding and supporting local partners, capacity can be enhanced through video and online courses for evaluators, and social media platforms can be used to help local evaluators connect and share good practice, learning, failures and materials. Again, these tools will only be useful if they are accessible and tailored to the context of evaluators.

11. Measure what matters

Results-based, data-based focuses can be biased towards the "countable" and leave out the complex-

ity and in-depth analysis made possible through collection of qualitative information. Many of the most common ICT tools used for M&E are designed to collect increasingly larger amounts of quantitative data. However efforts should be made to ensure that the M&E process also includes qualitative feedback where indicated.

Social media platforms can play a big role in engaging practitioners with evaluation results and in helping evaluators understand practitioner experiences.

12. Use and share M&E information effectively

Data visualizations during monitoring can be of enormous support in making decisions about program modifications and budget allocation, but it is important to ensure that data visualizations are not a goal. Rather, use of data to support decision-making is the important part. Real-time data and dashboards that allow program managers to keep up-to-date on progress are possible with new ICTs, yet these need to be accompanied by appropriate decision-making channels and authority levels. Sharing evaluation results can be of great value to organizations and the wider field of development, and social media platforms can play a big role in engaging practitioners with evaluation results and in helping evaluators understand practitioner experiences. It is important to think through and map out the different levels of data that will be collected – from mobile phone to crowdsourced or self-reported data, to an MIS, to a dashboard, to social media – and how existing data can be linked to other existing datasets. Without this, time and resources will be wasted collecting data that are similar, but not very useable by colleagues and partners.

References

Bamberger, M., Rugh, J., & Malory, L. (2012a). *Real World Evaluation: Working under budget, time, data and political constraints* (2nd Edition). Sage. pp 81-93.

Bamberger, M., Rugh, J., & Malory, L. (2012b). *Real World Evaluation: Working under budget, time, data and political constraints* (2nd Edition). Sage. pp 410-413.

Banga, K., Liesman, T., Meulensteen, A., & Wiemer, J. (2009). *Budgets, Batteries, and Barriers: PDA Implementation Issues for NGOs*.

Boyera, S., & Alonso, J. M. (2012). Mobile Data Collection in Africa. Available at: http://public.webfoundation.org/2013/data-collection/findings.pdf

Bruce, K. (2013). Use of Mobile Technology for Monitoring and Evaluation in International Health and Development Programs. Ph.D. Thesis.

Byrne, D. (2009) "Complex realist and configurational approaches to cases: a radical synthesis" in Byrne and Ragin (eds) 2009. The Sage Handbook of Case-Based Methods. Sage Publications.

Cabinet Office Behavioural Insights Team. (2012). *Applying behavioural insights to reduce fraud, error and debt*. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/60539/BIT_FraudErrorDebt_accessible.pdf

Cabinet Office Behavioural Insights Team. (2013a). *Applying behavioural insights to charitable giving*. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/203286/BIT_Charitable_Giving_Paper.pdf

Cabinet Office Behavioural Insights Team. (2013b). *Applying Behavioural Insights to Organ Donation: preliminary results from a randomised controlled trial*. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/267100/Applying_Behavioural_Insights_to_Organ_Donation.pdf

Chapelier, C., & Shah, A. (2013). Improving communication between humanitarian aid agencies and crisis-affected people: Lessons from the infoasaid project. *Network Paper*, (74). Available at: http://www.odihpn.org/hpn-resources/network-papers/improving-communication-between-aid-agencies-and-crisis-affected-people-lessons-from-the-infoasaid-project

Data-Pop Revised Concept Note, February 16, 2014 (unpublished document).

Davies, R. and Dart, J. (2005) The Most Significant Change (MSC) Technique. A Guide to Its Use. Available at: http://www.mande.co.uk/docs/MSCGuide.pdf

Demombynes, G. (2011). Collecting survey data via mobile phone in Southern Sudan. *Africa Can End Poverty*. Available at: http://blogs.worldbank.org/africacan/collecting-survey-data-via-mobile-phone-in-southern-sudan

Demombynes, G., Gubbins, P., & Romeo, A. (2013). Challenges and Opportunities of Mobile Phone-Based Data Collection Evidence from South Sudan. Available at: http://elibrary.worldbank.org/doi/pdf/10.1596/1813-9450-6321

Education Development Center. (2013). *Akazi Kanoze Youth Livelihoods Program: SMS-based system overview*.

Farmer, L., & Boots, M. (2013) Barriers and Solutions in using M4D: Connecting Directly to Citizens for Scalable Impact. VOTO Mobile Whitepaper. Available at: http://www.votomobile.org/files/M4D_Barriers_Whitepaper.pdf

Findley, A. M. G., Gleave, M. C., Morello, R. N., & Daniel, L. (2012). Title: Extrinsic , Intrinsic , and Social Incentives for Crowdsourcing Development Information in Uganda : A Field Experiment. Available at: http://www.aiddata.org/weceem_uploads/_ROOT/File/Crowdsourcing Field Experiment Uganda.24May2012.pdf

Forss, K; Marra, M, and Schwartz, R. (eds) 2011 Evaluating the complex: attribution, contribution and beyond. Transaction Press

Funnell, S. C., & Rogers, P. J. (2011). Purposeful program theory: Effective use of theories of change and logic models. San Francisco, CA: Jossey Bass.

Furubo, J.E., Rist, R., Speer, S.. (eds) 2013 Evaluation and Turbulent Times. Transaction Press.

Goldstein, M. (2012, July). Paper or Plastic? Part II: Approaching the survey revolution with caution. Impact Evaluations Blog. Available at: http://blogs.worldbank.org/impactevaluations/paper-or-plastic-part-ii-approaching-the-survey-revolution-with-caution

Gow, G., Quinn, K., & Barlott, T. (2013). Sexual Exploitation Outreach with Text Messaging: Introducing Project Backpage. Retrieved May 21, 2014, from http://www.frontlinesms.com/wp-content/uploads/2014/01/frontlinesms_casestudy_Sexual-Exploitation-Outreach-with-Text-Messaging.pdf

Groves, J (2013) Aiming for utility in "Systems-based evaluation": a research-based framework for practitioners. (PowerPoint presentation) Available at: http://www.slideshare.net/idsuk/paper-session-1-grove

Haushofer, J., & Neihaus, P. (2013). Optimizing Impact Evaluation for the Mobile Era. *Policy Design and Evaluation Lab*, 14–15. Available at: http://pdel.ucsd.edu/research/mobile-impact-evaluation.html

Heeks, R. (2009). The ICT4D 2.0 Manifesto: Where Next for ICTs and International Development? (Development Informatics Group, Ed.). Manchester, UK: Development Informatics. Available at: http://www.sed.manchester.ac.uk/idpm/research/publications/wp/di/documents/di_wp42.pdf

Insight Share (2012). *Participatory Video and Most Significant Change: How it works*. Available at: http://www.insightshare.org/sites/insightshare.org/files/Participatory%20Video%20and%20the%20Most%20Significant%20Change%20Technique%20-%20How%20it%20works.pdf

International Telecommunication Union (2013a). *The World in 2014. ICT Facts and Figures*. Available at: http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2014-e.pdf

International Telecommunication Union (2013b). *Measuring the Information Society 2013*. Geneva, Switzerland. Available at: http://www.itu.int/en/ITU-D/Statistics/Documents/publications/mis2013/MIS2013_without_Annex_4.pdf

Kane, M., & Trochim, W. M. K. (2007). *Concept mapping for planning and evaluation*. Thousand Oaks, CA: Sage Publications.

Kumar, S. (2002) Methods for Community Participation: a complete guide for practitioners. London. ITDG.

Kuruvilla, M. (2013) "Frontline SMS helps The Busara Center for Behavioral Economics cut 16 hours of work down to 30 minutes." Aug. 27. Available at: http://www.frontlinesms.com/2013/08/27/frontlinesms-helps-the-busara-center-for-behavioral-economics-cut-16-hours-of-work-down-to-30-minutes/

Lamb, J. (2013). *Community Mapping to Protect Forests in Cambodia*. Presentation at Interaction Forum.

Letouzé, E. (2013) *Could big data provide alternative measures of poverty and welfare?* Available at: http://www.developmentprogress.org/blog/2013/06/11/could-big-data-provide-alternative-measures-poverty-and-welfare

Letouzé, E. (2014) *Big data for development: Facts and figures*. Available at: http://www.scidev.net/global/data/feature/big-data-for-development-facts-and-figures.html

May, M. (2013, February 19). Crowdmapping the world we want. *Brac Blog*. Available at: http://blog.brac.net/2013/02/crowdmapping-the-world-we-want/

Meier, P. (2013). The art and science of delivery. *Voices on Society, Vol. 5*. Available at: http://voices.mckinseyonsociety.com/crisis-maps/

Muniz, S. (2013). Participatory video for M&E: supporting bottom-up learning. *ONTRAC Newsletter*, (55), 5. Available at: http://www.intrac.org/data/files/resources/775/ONTRAC-55-New-technologies-in-monitoring-and-evaluation.pdf

Plan (2013) GIS for Governance Program. Quarterly Report, October, shared by Plan Cameroon via email.

Pratihast, A. K., Herold, M., Avitabile, V., de Bruin, S., Bartholomeus, H., Souza, C. M., & Ribbe, L. (2012). Mobile devices for community-based REDD+ monitoring: a case study for Central Vietnam. *Sensors (Basel, Switzerland)*, *13*(1), 21–38. doi:10.3390/s130100021

Raftree, L. (2012a, August 9). 12 tips on using ICTs for social monitoring and accountability. *Wait... What?* Available at: http://lindaraftree.com/2012/08/09/tips-on-using-icts-for-social-monitoring-and-accountability/

Raftree, L. (2012b, September 12). 11 points on strengthening local capacity to use new ICTs for monitoring and evaluation. *Wait... What?* Available at: http://lindaraftree.com/2012/09/03/11-points-on-strengthening-local-capacity-to-use-new-icts-for-monitoring-and-evaluation/

Raftree, L. (2012c, October 12). 10 tips on using new ICTs for qualitative monitoring and evaluation. *Wait... What?* Available at: http://lindaraftree.com/2012/10/22/10-tips-on-using-new-icts-for-qualitative-monitoring-and-evaluation/

Raftree, L. (2013a, August 17). 12 lessons learned in ICTs for monitoring and accountability. ICT Works. Available at: http://www.ictworks.org/2012/08/17/12-lessons-learned-icts-monitoring-and-accountability/

Raftree, L. (2013b, April 17). Benefits, barriers and tips for ICT-enabled M&E. *Wait... What?* Available at: http://lindaraftree.com/2013/04/17/benefits-barriers-and-tips-for-ict-enabled-me/

Raftree, L. (2013c, March 12). ICTs and M&E at the South Asia Conclave Evaluators' Meeting. *Wait... What?* Available at: http://lindaraftree.com/2013/03/12/icts-and-me-at-the-south-asia-evaluators-conclave/

Rodin, J., & MacPherson, N. (2012). Shared Outcomes. *Advancing Evaluation Practices in Philanthropy*, (Stanford Social Innovation Review Special Supplement), 12–15. Available at: http://www.aspeninstitute.org/sites/default/files/content/docs/pubs/PSI-SSIR-Advancing-Eval-Practices-Philanthropy.pdf

Ross, D. (2012, August). Harnessing Sport and ICTs against HIV/AIDS – GOAL Trial of Sport-based HIV Prevention in South African Schools run by LIDC members from LSHTM and IOE. *Blog of the London International Development Centre*. Available at: http://lidcblog.wordpress.com/2012/08/08/harnessing-sport-and-icts-against-hiv-aids-goal-trial-of-sport-based-hiv-prevention-in-south-african-schools-run-by-lidc-members-from-lshtm-and-ioe/

Stahl, B., Heersmink, R., Goujon, P., Flick, C., van den Hoven, J., Kutoma, W., ... Radar, M. (2010). Identifying the Ethics of Emerging Information and Communication Technologies: An Essay on Issues, Concepts and Method. *International Journal of Technoethics*, *1*(4), 20–38. Available at: http://www.tech.dmu.ac.uk/~bstahl/index_html_files/2010_etica_methodology_IJT.pdf

Thakkar, M., Floretta, J., Dhar, D., Wilmink, N., Sen, S., Keleher, N., Shaughnessy, L. (n.d.). *Mobile-Based Technology for Monitoring and Evaluation*. Available at: http://www.theclearinitiative.org/mobile-based-tech.pdf

The Rockefeller Foundation. (n.d.). *Six Factors that Boost an Organization's Capacity to Innovate*. Available at: http://www.aspeninstitute.org/sites/default/files/content/docs/pubs/PSI-SSIR-Advancing-Eval-Practices-Philanthropy.pdf

The Rockefeller Foundation (2012) Storytelling Peer Review Meeting. Notes and Summary. Internal Document. The Rockefeller Foundation, September 24.

Tonea, D. (2013). ICTs: an effective way to improve beneficiary complaint mechanisms? *ONTRAC Newsletter*, (55), 6. Available at: http://www.intrac.org/data/files/resources/775/ONTRAC-55-New-technologies-in-monitoring-and-evaluation.pdf

Trigg, C. (2013). Practical challenges and advantages of using digital devices for survey work. *ONTRAC Newsletter*, (55), 4. Available at: http://www.intrac.org/data/files/resources/775/ONTRAC-55-New-technologies-in-monitoring-and-evaluation.pdf

UN Global Pulse. (2012). *Big Data for Development: Challenges & Opportunities*. Available at: http://www.unglobalpulse.org/sites/default/files/BigDataforDevelopment-UNGlobalPulseJune2012.pdf

UNICEF. (2012, March 14). U-Report application revolutionizes social mobilization, empowering Ugandan youth. *UNICEF website*. Available at: http://www.unicef.org/infobycountry/uganda_62001.html

Underwood, H., Sterling, S. R., & Bennett, J. K. (2012). The PartoPen in Training and Clinical Use: Two Preliminary Studies in Kenya. Available at: http://homes.cs.washington.edu/~rohitc/HealthINF_Underwood_CameraReady.pdf

Wakunuma, K., (2010) Doctoral dissertation.

Walker Hudson, L. (2013). Mobile Phones in M&E: towards sustainable and democratic practice. *ONTRAC Newsletter*, (55), 7. Available at: http://www.intrac.org/data/files/resources/775/ONTRAC-55-New-technologies-in-monitoring-and-evaluation.pdf

Yu, P, de Courten, M., Pan, E., Galea, G., Pryor, J. The development and evaluation of a PDA-based method for public health surveillance data collection in developing countries. Int J Med Inform 2009;78:532-42.

Financial support provided by
THE ROCKEFELLER FOUNDATION

www.rockefellerfoundation.org