Independent Evaluation at MCC

An Evolving Practice 18 Years On

Shreena Patel and Cindy Sobieski

January 2023

White Paper:
Independent Evaluation at MCC
An Evolving Practice 18 Years On

Shreena Patel, Senior Director of Evaulation

Cindy Sobieski, Director of M&E

January 2023

TABLE OF CONTENTS

The Millennium Challenge Corporation's (MCC) Monitoring and Evaluation (M&E) approach is based on three core principles: **accountability**, **learning**, and **transparency**. This paper examines how MCC M&E applies these core principles to its evaluation practice, striving for high-quality results measurement and use. The paper first reviews the foundational principles of MCC's evaluation approach. It then describes the practices MCC M&E employs to embody these principles and explains how they evolved over time as the agency learned from experience. Finally, the paper presents a refined definition of high-quality evaluation at MCC. Going forward, MCC will continue to strengthen its evaluation practice, always striving to achieve the highest quality evaluations possible, given practical constraints.

EXECUTIVE SUMMARY

FOUNDATIONAL PRINCIPLES

The Millennium Challenge Corporation (MCC) is an evidence-based US government foreign assistance agency whose model is based on principles of good governance, country ownership, a focus on results, and transparency. To operationalize the "focus on results," project teams design and manage for results, while the Monitoring and Evaluation (M&E) team measures results to produce evidence on development effectiveness. Using the core principles as a foundation, MCC's M&E practice is guided by the following:

✓ **Accountability**: Measure the extent to which each MCC-funded program achieved what it intended to.

📖 **Learning**: Identify why an investment did or did not achieve its stated objective to inform adaptation in current and future programming.

🔍 **Transparency**: Publicly release information about the measured results of MCC-funded programs.

MCC's M&E policy has always been to independently evaluate every project, and each country program includes a specific M&E budget for this purpose. MCC and partner country M&E staff are also integrated into program teams early in program design to facilitate a shared understanding about the results that projects are seeking to achieve and that will need to be measured before, during, and after implementation.

Over the last 18 years, MCC's evaluation practice consistently reflected on its work and adapted to improve. This paper builds on prior publications of M&E learning to describe the key drivers of this evolution, the systems for managing evaluation quality, and a refined definition of high-quality evaluation to guide MCC's evaluation practice going forward. It concludes with questions about how MCC evaluation may adapt in new ways in the future.

AN EVOLVING PRACTICE

MCC has been consistent in requiring a rigorous, independent evaluation for all projects, but has reflected on experience and adapted its internal processes, methods, and approaches over time.

The evolution of MCC's evaluation practice, since its start in 2004, can be organized into the three phases.

Piloting Foundational Principles (2004-2012)

During the early years of MCC, evaluations were characterized by a priority on impact evaluation, a focus on measuring household incomes as the key result of all projects, and a commitment to

the independence of evaluators. An initial set of completed MCC program evaluations, released in 2012, provided a valuable opportunity to reflect on how these features fared in practice and offered *several lessons* that shaped the *next phase* of MCC evaluation, including:

📖 A project logic, or theory of change, must serve as the foundation for evaluation design

📖 Impact evaluation may not be appropriate in all cases

📖 Evaluation plans need project team buy-in and oversight

Setting up a System (2012-2017)

As MCC's program portfolio grew and matured, the number of evaluations M&E staff were managing grew exponentially. This required a robust system to manage quality and consistency across the evaluation portfolio. The M&E team launched a new standardized evaluation management system in 2013, which consisted of written guidance to establish evaluation quality standards, structures to guide MCC staff and contractors in working toward these quality standards, and tools to promote use of evaluation evidence.

To define quality standards, MCC developed the Evaluation Management Guidance. It laid out MCC's evaluation management approach in terms of (i) roles and responsibilities, (ii) the formal review system for evaluation materials and reports, and (iii) the milestones and risk factors. Significant revisions to the Guidance were undertaken in 2017 and 2020 to provide additional direction on how to structure evaluations, better track data on the evaluation portfolio, and optimize evaluation resources. MCC also developed a standard evaluation scope of work and outlines for each type of evaluation report to define what is expected from an MCC-funded evaluation deliverable.

In line with the agency's 2012 commitment to provide open data to the public, MCC defined standards for ethical data management and publication in the Transparent, Reproducible, and Ethical Data and Documentation (TREDD) Guidelines. MCC also established a Disclosure Review Board to consider risks before publishing any data for public use or restricted access.

Building off these quality standards, MCC developed various management structures to ensure these standards are put into practice as effectively as possible, most notably the Evaluation Management Committee. It is convened at the start of each evaluation and is comprised of sector and cross-cutting technical staff working on a given project. It aims to coordinate across project design, implementation, and evaluation and build agency buy-in for attainment and use of evaluation results. Committees follow explicit guidance about the scope of their feedback when reviewing evaluation materials to maintain the independence of the evaluation.

Advancing Evidence Use (2017-present)

By 2017, MCC's evidence generation system had reached a steady state, with a large number of completed evaluations to drawn from. This

prompted a shift in the M&E team's focus to improving evidence use. The following products and tools were revised or newly developed during

this phase to better present evaluation results and lessons and make them more easily findable and searchable. These innovations intend to help evaluation users quickly access exactly the information they need.

Evaluation Briefs – four-page summaries of each evaluation's results, lessons, and methods, designed to serve various users' needs. They are standardized and visual summaries as well as entry points to more detailed evaluation reports and materials.

MCC Learning – detailed lessons for future programming and evaluation identified by the Evaluation Management Committees, as motivated by the results of each evaluation.

Evidence Platform – public website sharing evaluation reports, questionnaires, data, and other documentation with enhanced search, filtering, and sharing functions.

Evaluation Dissemination Workshops – presentations convened at both MCC headquarters and in partner countries to facilitate engagement with evaluation results and learning.

Sector Results and Learning Pages – public websites that compile sector-specific evidence, by mapping programs and common interventions, aggregating sector data on key performance indicators, and presenting a curated set of evaluation results along with a filterable database of all sector program lessons identified in the MCC Learning.
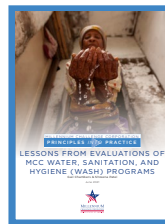
Evidence Synthesis Papers – publications reflecting on a critical mass of implementation experiences and evaluation results on a topic, drawing broader lessons to improve future sector programming and M&E.

HIGH-QUALITY EVALUATION: REFINING THE DEFINITION

The core mission of M&E staff has been to produce as high-quality an evaluation as possible, however, what constitutes high-quality was articulated piecemeal, leading to different points of focus across evaluations. MCC now defines a quality evaluation as one that:

- answers two summative evaluation questions regarding project effectiveness

- is conducted by parties independent of project design and implementation

- uses rigorous and ethical methods to produce unbiased analysis that is meaningful for decision-makers

This definition serves as a guide or a goal to aim for and will be refined over time. Key details include:

- To fulfill MCC's commitment to accountability, each evaluation should clearly answer summative questions about the scope and quality of project implementation, whether the project achieved its stated objective in the timeframe and magnitude expected, and why or why not.

- Evaluation rigor does not solely relate to methodology and depends as well on factors such as the clarity of the logic being evaluated, the feasibility of the evaluation design, and the quality of data collection materials and methods and the analysis.

- Ethical evaluations must adhere to promises of respondent confidentiality to reduce risk of direct harm to data providers.

- Relying on independent evaluators and upholding public transparency about evaluation results and the data and analysis underlying them, are critical to producing unbiased assessments.

- To meaningfully support accountability and learning to achieve greater development impact, an evaluation should be relevant, respected, trusted, timely, and accessible.

MCC EVALUATION OVER THE NEXT DECADE

MCC's unique evidenced-based model, underpinned by its core principles, provides the foundation for a high-quality evaluation practice to support achievement of MCC's mission. Considering the decade ahead, the following questions may guide further evolution within MCC's evaluation practice:

- What is the right level for accountability (e.g. country program, project, or each funded intervention)?

- Is the project objective the right anchor for evaluations?

- Is a focus on accountability adding value for MCC?

- Can MCC's model be optimized for learning?

- Does MCC's evaluation practice strike the right balance between accountability and learning?

These questions are all substantial and likely relevant to other international development organizations. They also provide an exciting opportunity for MCC to lead on how best to use evaluation, accountability, and learning to improve international development practice. With these and other questions in mind, MCC M&E has and will continue to evolve to ensure that the evidence it produces is high-quality and is being put to its best use.

1 FOUNDATIONAL PRINCIPLES

The Millennium Challenge Corporation (MCC) is an independent US foreign assistance agency with a mission to reduce poverty through economic growth. It is an evidence-based institution that executes policy and investment decisions using available empirical evidence, development theory, and international best practices. Since the passage of the Millennium Challenge Act of 2003, which founded the agency, MCC has worked with 49 low- and lower-middle income partner countries, providing time-limited grants through 68 programs.

MCC's model is based on a set of core principles essential for effective development assistance — good governance, country ownership, a focus on results, and transparency. [1] To operationalize the "focus on results," project teams design and manage for results, while the Monitoring and Evaluation (M&E) team measures results. This produces evidence on the effectiveness of MCC's investments to inform future agency decisions. Drawing from this agency-wide foundation, MCC M&E seeks to operationalize three principles: **accountability, learning,** and **transparency**.

---

1   MCC core principles are referenced here: *https://www.mcc.gov/our-impact/principles-into-practice*

Accountability

As a publicly financed entity, MCC strives to hold itself and its partner countries mutually accountable to taxpayers and the groups expected to benefit from MCC programs. This is accomplished by determining the extent to which a program achieved what it intended to. MCC's authorizing legislation requires program agreements, called compacts, to state "the specific objectives that the country and the United States expect to achieve during the term of the compact" and "regular benchmarks to measure, where appropriate, progress toward achieving such objectives".[2][3]

Learning

MCC strives to learn why an investment did or did not achieve what it intended to (i.e., its stated objective). These findings are designed to inform decisions regarding the design, implementation, analysis, and measurement of current and future interventions.

Transparency

MCC is committed to publicly releasing as much information as is feasible and cost-effective regarding the monitoring and evaluation of its programs. MCC's website is routinely updated with the most recent monitoring information. Partner country entities, known as Millennium Challenge Accounts (MCAs), are encouraged to do the same on their respective websites. For each evaluation, MCC posts all associated results, reports, data collection materials, lessons learned and de-identified data.

MCC's evaluation practice also aligns with broader US government legislation including the Foreign Aid Transparency and Accountability Act of 2016 and the Foundations for Evidence-Based Policymaking Act of 2018. It takes dedicated M&E resources to implement evaluation in line with these principles and legislation. MCC's M&E policy has always been to evaluate every project, and include a specific M&E budget for every program.[4] MCC and partner country M&E staff are integrated into program teams early in program design.

This practice facilitates a shared understanding of expected results between program designers and M&E staff . It also allows MCC and its partners to set up monitoring systems prior to program implementation and prospectively design independent evaluations that extend beyond the program period to measure results when they were expected to manifest.
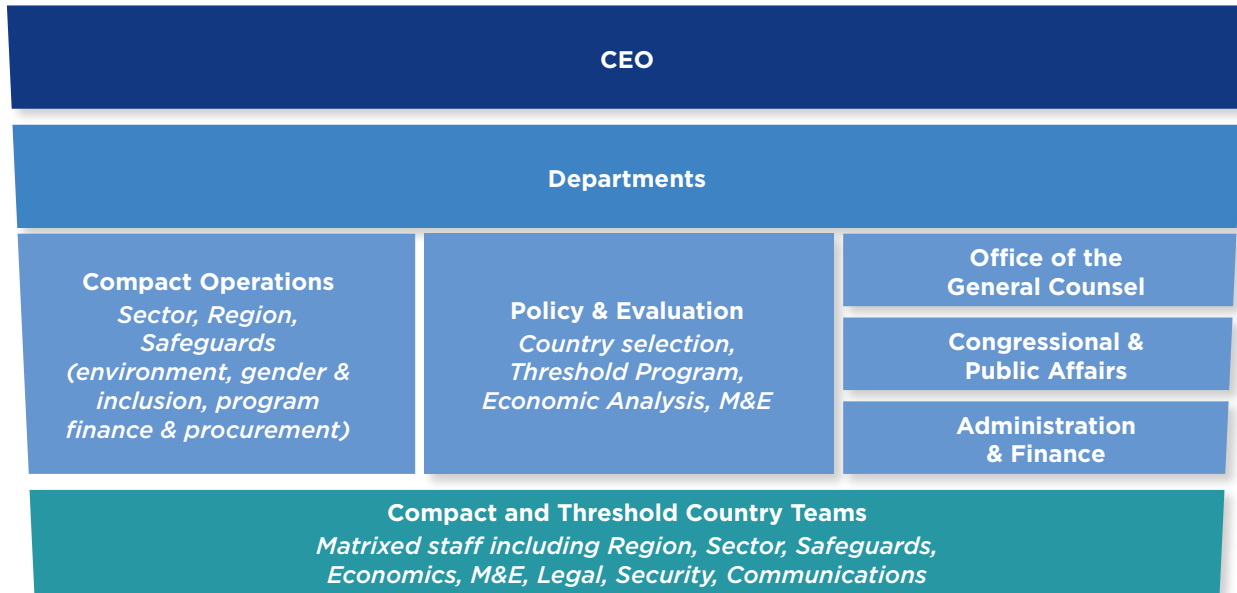
---

2   Project Objectives are the primary outcome that a Project intends to achieve to be considered successful; they are defined in Section 1.2 of the program agreement (Compact or Threshold).

3   MCC's authorizing legislation is available here: *https://assets.mcc.gov/content/uploads/mca-legislation-2018-amendments.pdf*

4   On average, 3% of a project's budget are spent on M&E.

Over the last 18 years, MCC's evaluation practice consistently reflected on its work and adapted to improve. M&E staff learned from the experience of managing a growing evaluation portfolio, responded to MCC's evolving program and evidence needs, and engaged in industry best practices. This paper builds on prior M&E team lessons to describe key drivers of this evolution, the systems for managing evaluation quality, and the refining of the definition of high-quality evaluation used to guide MCC's evaluation practice. [5] It concludes with questions about possible future MCC M&E adaptations.

CEO

Departments

Compact Operations	Policy & Evaluation	Office of the General Counsel
Sector, Region, Safeguards (environment, gender & inclusion, program finance & procurement)	*Country selection, Threshold Program, Economic Analysis, M&E*	Congressional & Public Affairs
		Administration & Finance

Compact and Threshold Country Teams
Matrixed staff including Region, Sector, Safeguards, Economics, M&E, Legal, Security, Communications

---

5 Prior evaluation learning papers available here: *https://www.mcc.gov/content/uploads/2017/05/paper-2012001116901-princi_ples-impact-evaluations.pdf* and *https://www.tandfonline.com/doi/full/10.1080/19439342.2014.975424*

2   AN EVOLVING PRACTICE

When applying its principles, MCC has consistently required a rigorous, independent evaluation for all projects. However, some aspects of the agency's evaluation approach have evolved. As the evaluation portfolio grew in size and complexity, MCC reflected on experience and adapted internal processes, methods, and approaches to both monitoring and evaluation. This section provides a brief overview of the most significant changes since MCC was established in 2004.

2.1   PILOTING FOUNDATIONAL PRINCIPLES (2004-2012)

At the outset, MCC established an evidence-based approach to implement the agency's commitment to results, which guided the focus and structure of early MCC evaluations. MCC M&E had not yet articulated its definitions of the core principles of accountability, learning, and transparency. During those early years , evaluations were characterized by three key features referenced in MCC's first M&E Policy (2009):

✔ A focus on impact evaluation as a primary evaluation methodology

✔ A focus on household income as the primary result to be measured by the evaluation

✔ A commitment to independent evaluation

MCC and the broader development community recognized the need to build more evidence on the causal impact of development interventions

through impact evaluations.[6] This initial emphasis was reflected in the 2009 M&E Policy, which prioritized the use of rigorous impact evaluations as the key set of methods for measuring MCC's results.[7]

Impact vs. Performance Evaluations

Impact evaluations (IE) are designed to measure changes in outcomes that can be attributed to the MCC investment using a valid counterfactual.

Performance evaluations (PE) are designed to measure changes in outcomes and assess the contribution of MCC investments to changes in those outcomes when it is not possible to identify a valid counterfactual.

Evaluations also focused on measuring program impact on household income. In particular, the policy at the time required an explicit link between a project's M&E Plan - which defines what will be measured by the evaluation - and its economic analysis. The economic analysis includes a cost-benefit analysis (CBA) to assess project proposals, model expected results, and estimate the economic rate of return (ERR). The methodology behind MCC's early CBAs focused on household income increases as the key result, leading to evaluations that focused on measuring incomes. However, the CBAs did not make the pathway to income gains clear. [8] With no theory of change or intermediate outcomes articulated, measuring income gains was challenging and sometimes unrealistic.
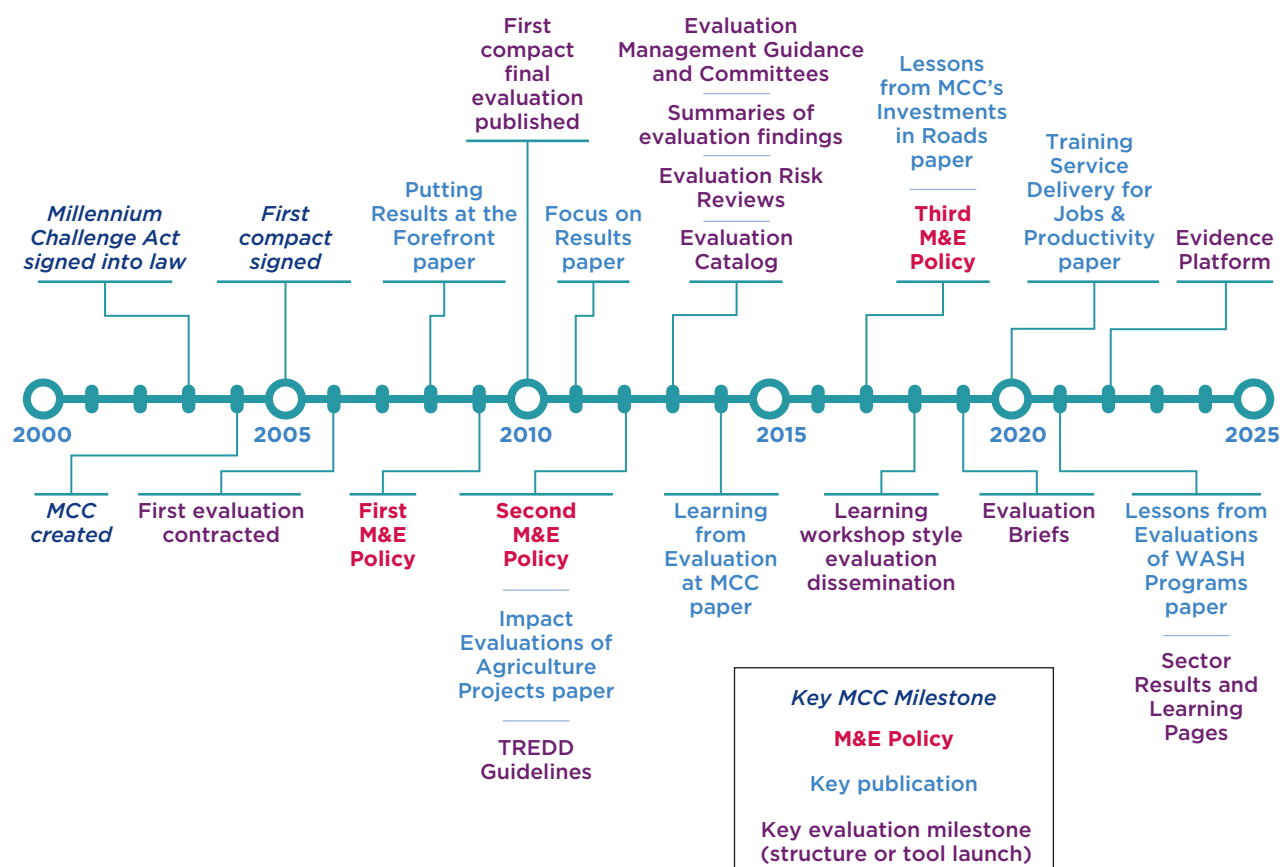
The third element reflected in the 2009 Policy aligned with principles for evaluation of development assistance established in 1991 by the Organization for Economic Co-operation and Development (OECD) Development Assistance Committee (DAC). It focused on the necessary impartiality and independence of the evaluation function and prohibited evaluators from participating in decision-making related to the design or implementation of the project being evaluated to ensure they had no stake in the success or failure of the project.[9]

---

6   Movement toward impact evaluation referenced here: *https://www.cgdev.org/publication/when-will-we-ever-learn-improving-lives-through-impact-evaluation*

7   MCC "Focus on Results" paper referencing emphasis on impact evaluation available here: *https://www.mcc.gov/content/uploads/2017/05/paper-2011001052001-principles-results.pdf*

8   MCC "Aid Effectiveness: Putting Results at the Forefront" paper referencing focus on household incomes available here: *https://www.mcc.gov/content/uploads/2017/05/mcc-112008-paper-results.pdf*

9   DAC guidelines referencing evaluation independence available here: *http://www.oecd.org/development/evaluation/2755284.pdf*

Timeline milestones (top to bottom of timeline):

Key publications (above timeline):
- Millennium Challenge Act signed into law
- First compact signed
- Putting Results at the Forefront paper
- First compact final evaluation published
- Focus on Results paper
- Evaluation Management Guidance and Committees
- Summaries of evaluation findings
- Evaluation Risk Reviews
- Evaluation Catalog
- Lessons from MCC's Investments in Roads paper
- Third M&E Policy
- Training Service Delivery for Jobs & Productivity paper
- Evidence Platform

Timeline years: 2000 — 2005 — 2010 — 2015 — 2020 — 2025

Milestones (below timeline):
- MCC created
- First evaluation contracted
- First M&E Policy
- Second M&E Policy
- Impact Evaluations of Agriculture Projects paper
- TREDD Guidelines
- Learning from Evaluation at MCC paper
- Learning workshop style evaluation dissemination
- Evaluation Briefs
- Lessons from Evaluations of WASH Programs paper
- Sector Results and Learning Pages

Legend box:
Key MCC Milestone
M&E Policy
Key publication
Key evaluation milestone (structure or tool launch)

In 2012, MCC released an initial set of compact program evaluations focused on agriculture, which provided an extremely valuable opportunity to reflect on MCC's early emphasis on rigorous impact evaluations in practice. Drawing on these agricultural evaluations and other early evaluations in process, MCC found several challenges with the implemented approach, including:

- Maintaining valid impact evaluation designs for complex investments that consisted of various interventions on different timelines.

- No guide for evaluation data collection, as the causal chain by which projects aimed to achieve income gains was not articulated in the project design or economic analysis.

- Difficulty in detecting significant changes in income, a challenging result to achieve and a difficult one to measure accurately.

- Balancing the need to keep evaluators independent from project decision-making with the need for evaluators to understand details of program design and implementation.

Reflecting on the challenges of early evaluations resulted in lessons that shaped the next phase of MCC evaluation practice, including:[10]

📖 Structuring evaluations around a well-defined project logic.

📖 Applying impact evaluation approaches only when feasible and appropriate, such as when project design and implementation are conducive.

📖 Coordinating evaluation plans with evaluators and the full project team at MCC and Millennium Challenge Account (MCA) to ensure evaluations are properly designed and implemented.

The full set of lessons is discussed in the MCC Principles into Practice paper "*Impact Evaluations of Agriculture Projects*" (2012). Additional reflections on this early stage of MCC's evaluation practice as well as the beginning of the next stage are available in "*Learning from evaluation at the Millennium Challenge Corporation*" (2014).
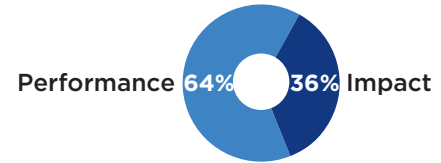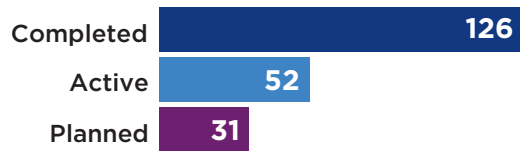

2.2 SETTING UP A SYSTEM (2012-2017)

This phase reflected a strong response to the lessons from early impact evaluations and the fast-growing evaluation portfolio. As MCC's program portfolio grew and matured, the number of evaluations M&E staff managed more than doubled between 2008 and 2011 - from 41 to 85 evaluations. By 2018 it more than doubled again to 195.[11] This

growth required a robust system to manage quality and consistency across the evaluation portfolio, as described below. Once the management system had been designed, it took time and dedicated oversight to become fully integrated. By 2017, the management system reached a steady state and it remains in place today.
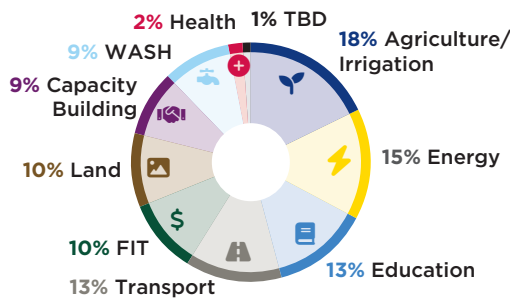
---

10   Most, though not all, of the compact evaluations released around 2012 were impact evaluations of agriculture projects. This set of agriculture impact evaluations strongly shaped MCC's evaluation practice at the time. Evaluations published prior to 2012 were mostly of MCC Threshold Programs implemented by USAID.

11   The number of evaluations grew exponentially at this time in part because M&E management portfolio reviews revealed that early MCC projects had included multiple components with distinct logics (and sometimes distinct populations and geographies), and that the evaluations of those projects had frequently prioritized evaluating components that could be randomized while leaving other large investments un-evaluated. The M&E team began filling in these gaps with evaluations tailored to each distinct logic within a project, which resulted in multiple new evaluations per project. The team continued to prospectively scope and commission separate evaluations for each distinct logic within a project until the revised guidance to focus on the project objective discussed later in this paper was developed.

State of the MCC Evaluation Portfolio - OCTOBER 2022

Completed **126**
Active **52**
Planned **31**

Performance **64%** **36%** Impact

168	**59**	**44**	**123**	**75**
Evaluation design reports	Baseline reports	Interim reports	Final reports	Data packages

Evaluations by Sector

2% Health 1% TBD
9% WASH 18% Agriculture/Irrigation
9% Capacity Building
10% Land 15% Energy
10% FIT
13% Transport 13% Education

$325 MILLION Estimated cost of contracted and completed evaluations

$10+ BILLION in MCC investment under evaluation

In building out the evaluation management system, the M&E team faced various decisions regarding scope and focus. To guide this decision-making, the M&E team better articulated and operationalized its definitions of accountability, learning, and transparency as they pertain to evaluation. In direct response to the early impact evaluation lessons, the 2012 M&E Policy codified the use of performance evaluation methods as an acceptable means of evaluating projects when impact evaluation methods are not feasible or appropriate. MCC also recognized the importance of detailing how proposed investments were expected to deliver meaningful results for targeted populations, as modeled in the cost-benefit analysis (CBA). As such, the agency required descriptions of the theory of change, or project logic, in proposed investment materials. The agency's M&E staff used these descriptions as the foundation for the investment's evaluation, resulting in evaluations designed to test an explicit and well-elaborated set of hypotheses, with targets and timelines for specific results. **Accountability** would be achieved by designing evaluations to validate the outputs and outcomes in this pre-determined project logic, and **learning** would come

from measuring not only the final outcome, but also the intermediate levels of the project logic. This shift represented a move away from centering MCC evaluations on measuring impacts on incomes.

In 2013 the M&E team developed a new standardized evaluation management system responding once again to early impact evaluation lessons. This system consisted of written guidance to establish evaluation quality standards and structures to guide the work of MCC staff and contractors toward the quality standards. These components were eventually codified in the 2017 M&E Policy. They can be grouped into three broad stages of the evaluation lifecycle:

📖 Defining evaluation purpose and scope

📖 Managing evaluation design and implementation

📖 Using evaluation-produced evidence and learning

The components are detailed throughout the next three sections.

2.2.1 Defining Quality Standards

The following guidance materials specify quality and completion standards for all aspects of the evaluation lifecycle. They ensure that the M&E staff overseeing the evaluation and the independent evaluators conducting it have a shared understanding of the expectations for all MCC evaluation components.

Evaluation Management Guidance[12]
In 2013, the M&E team established the *Evaluation Management Guidance* (the Guidance), to govern the structure and management of MCC's independent evaluations. It defined MCC evaluation types and methodologies, enabling tracking of evaluation characteristics across the portfolio. Most notably, it established the interdepartmental Evaluation Management Committee (the Committee), to ensure shared ownership and evaluation quality, which is discussed in detail below. [13] Finally, it laid out MCC's evaluation management approach in terms of:

- Roles and responsibilities
- The formal review system for evaluation materials and reports
- Milestones and risk factors

The formal review process placed equal emphasis on MCC and partner country stakeholders, including MCA staff, project implementers, and relevant government offices, so both parties engage in evaluation decision-making and have opportunities to review all data collection materials and reports prior to finalization.

The Guidance is a living document that is frequently updated to reflect M&E lessons and evolving approaches. Significant revisions to the Guidance were undertaken in 2017 and 2020 to provide additional direction on structuring evaluations, including reorienting them around the assessment of project objectives, better data tracking on the evaluation portfolio, and optimizing evaluation resources.

To support consistent application of the Guidance, the M&E team developed a standard evaluation scope of work and outlines for evaluation reports. The standardized scope of work was launched in 2013 to define the tasks and deliverables each evaluation should produce, in line with applicable MCC policies and guidance. For each evaluation, MCC and country partners establish the evaluation scope by articulating the evaluation questions. Learning from MCC's early evaluations, the scope of work does not include any text prescribing evaluation design elements beyond the evaluation type (impact or performance) to allow evaluators the freedom to propose innovative methods. While MCC originally allowed the design of evaluations to be completely independent, this practice misaligned incentives between evaluators and MCC regarding the level of complexity and cost of proposed evaluation designs. It also invited scope creep to research aims beyond project evaluation. The Committee now engages more closely on design decisions to better manage costs and benefits, however MCC strongly maintains the line of independence when it comes to evaluation analysis. The scope of work has been revised periodically to reflect evolutions in the M&E team's standards and approaches, including to incorporate standards for data de-identification and dissemination.

---

12   *https://www.mcc.gov/resources/doc/guidance-evaluation-management*

13   Evaluation types: Impact, Performance, or Both. Evaluation methodologies: **Impact**: (i) Random Assignment; (ii) Continuous Treatment; (iii) Difference in Differences only; (iv) Matching only; (v) Difference in Differences with Matching (vi) Interrupted Time Series; (vii) Regression Discontinuity, and (vii) Other Impact (i.e., other quasi-experimental methods). **Performance**: (i) Pre-Post; (ii) Ex-Post thematic analysis; (iii) Pre-Post with Comparison Group; (iv) Modeling; and (v) Other Performance.

Refining the meaning of accountability in MCC evaluation

✓ From MCC's beginnings, the principle of **accountability**, as reflected in MCC's founding legislation requirement to measure progress toward achieving objectives, has guided independent evaluations. However, exactly how the M&E team scopes evaluations for **accountability** has evolved.

For example, the first version of the Guidance (2013) framed the key evaluation question as:

- *Were there any changes in key outcomes, such as income, for program participants that are attributable to MCC's investment?*

The revision (2017) asked:

- *Were there any changes in key outcomes for program participants that may be attributable to MCC's investment?*

By 2020, M&E staff had reviewed hundreds of design, baseline, interim, and final evaluation reports, and found that stakeholders frequently looked to evaluations as research studies exploring various theories rather than as evaluations of a defined project. For example, many evaluations investigated results beyond the theory of change. These evaluations struggled to measure such a large set of results, each requiring different samples and methods. M&E staff recognized that each evaluation faces tradeoffs to balance its quality, costs, and benefits and understood that the team must make hard decisions regarding evaluation scope up front.

As such, the M&E team adjusted the standard evaluation questions in the 2020 Guidance to make the purpose and scope of evaluations clear, making explicit reference to the project objective:

- *Did the program achieve its targeted outcomes, particularly its stated objective, in the timeframe and magnitude expected? Why or why not?*

The project objective is the primary outcome that a project intends to achieve to be considered successful, as defined in Section 1.2 of the program agreement (Compact or Threshold). This clarified the definition of **accountability** as a rigorous assessment of whether the project achieved its stated objective as expected and highlighted the project objective as the cornerstone of each evaluation.

While the focus on the objective helped, M&E staff still grappled with a lack of clarity around what constituted a complete project evaluation. Specifically, what results must be measured to consider a project evaluated? MCC projects may contain theories of change leading to results unrelated to the objective statement, in addition to the theory leading to the objective itself. Analyzing all these standalone theories spread resources thin, resulting in multiple evaluations per project and unfocused evaluation reports that did not provide a clear answer to whether MCC had achieved what it set out to (**accountability**).

To optimize the use of resources, produce clear and actionable evidence, and stay true to its foundational principles, the M&E team shifted again to focus on the project objective and the logic leading to it. Thus, measuring and analyzing this theory is sufficient to consider the project "evaluated" and this is reflected in the definition of quality evaluations presented in Section 3.

In line with other movements for standardized evaluation reporting, the M&E team complemented the standard scope of work with established standard reporting guidelines for the Evaluation Design Report, Baseline Report, and Interim/Final Reports. [14]These materials define what is expected from an MCC-funded evaluation deliverable, increasing the accuracy of evaluation budget estimates and the efficiency of reviews of evaluation products. Standardized products ensure key details are easy to find when needed for MCC decision-making and to facilitate meta-analysis.

Transparent, Reproducible, and Ethical Data and Documentation (TREDD) Guidelines[15]

In 2012, MCC operationalized a commitment to provide open data to the public in line with MCC's principle of **transparency** and broader US government initiatives. As discussed in the MCC Principles Into Practice paper on *transparency*, this commitment required careful thought about ethical data management and publication, particularly with primary evaluation data containing personally identifiable information. [16] MCC established a Disclosure Review Board to consider these risks. MCC evaluators began preparing data from all independent evaluations for publication and the M&E team leveraged this practice to further promote the rigor and credibility of independent evaluations. Evaluators document their data sets and analysis code such that results can be replicated by external data users.

Through this experience, M&E staff learned that striking the right balance between **transparency**, reproducibility of results, and ethical obligations requires careful oversight. For example, some de-identification decisions result in the loss of reproducibility. MCC's original Transparent, Reproducible, and Ethical Data and Documentation (TREDD) guidance was established in 2012 to define best practices for MCC staff and contractors while collecting, managing, and publishing data collected from human subjects. Significant revisions to TREDD were led in 2017 and 2020 to reflect changes to the Federal Policy for the Protection of Human Subjects and to organize TREDD guidance around the evaluation lifecycle.[17]

---

14   Movement toward standardized reporting referenced here: *https://www.equator-network.org/*

15   *https://www.mcc.gov/resources/doc/guidance-mcc-guidelines-tredd*. Originally titled MCC Microdata Management Guidelines.

16   MCC "Transparency" paper available here: *https://assets.mcc.gov/content/uploads/2017/05/paper-2015001163301-principles-transparency.pdf*

17   Federal policy on the protection of human subjects available here: *Revised Common Rule | HHS.gov*

2.2.2 Managing toward quality

Building off the quality standards described above, the M&E team developed various management structures to ensure these standards are put into practice as effectively as possible. They remain in place today but have been updated as needed.

Evaluation Management Committees
The M&E team established the Evaluation Management Committee to respond to the need for better coordination across project design, implementation, and evaluation and to build agency buy-in for attainment and use of evaluation results. It consists of sector and technical staff working on a given project, reflecting relevant practice groups across the agency. Committees are established for each evaluation to guide evaluation decisions and ensure evaluations reflect input and expertise from key stakeholders across MCC. Committee member roles include the following:

- Project leads review evaluation materials to ensure that results and technical terms are defined correctly and reflect the project and its local context. Their engagement mitigates the distance between evaluators and implementers, which hampered MCC's early evaluations.
- Economists review materials to ensure that the evaluation's measures align with the results defined in the economic analysis.
- M&E staff bring expertise in evaluation and data collection methods and have the primary responsibility for managing evaluation quality, Committee reviews, and decision-making.

This level of engagement requires a significant time commitment from Committee members. It also requires M&E staff to play a coordination role, both ensuring members' input and advancing the work. However, MCC continues to support this engagement across the agency as part of its focus on results.

The Committee is the M&E team's primary mechanism for ensuring the evaluation's expected benefits are consistently weighed against the expected risks and costs, and any necessary course corrections – such as changes to scope, design, and budget – are made to maintain balance between evaluation costs and expected **accountability** and **learning** benefits.

To maintain the independence of the evaluator's analysis, the Committee follows explicit guidance about the scope of their feedback when reviewing evaluation materials. To avoid undue influence on results, this guidance becomes stricter for interim or final report reviews, which assess program effectiveness. To maintain **transparency** and evaluator independence, Committee members understand that evaluators are not required to incorporate all reviewer feedback into evaluation reports, and Committee comments and the evaluator's responses will be made public. Names are removed to encourage candid feedback.

In line with MCC's principles of country ownership and **accountability**, MCA M&E staff engage closely in the design and implementation of the evaluation, often directly contracting and overseeing data collection in-country. Either through their own Committee or a similar informal structure tailored to the MCA's organizational structure, they ensure other MCA staff and relevant partner government stakeholders buy into the evaluation scope and design and have the opportunity to review and comment on all evaluation materials. MCA reviews further ensure that the evaluation work reflects realities on the ground. Partner country stakeholder comments are made public along with MCC's. Evaluators must carefully consider all feedback from MCC and partner countries and both parties have the opportunity to review

evaluator responses and edits to ensure feed-back was adequately considered before the data

collection materials and evaluation reports are finalized.

Evaluation Risk Reviews

To catch some of the challenges that plagued MCC's early evaluations before getting too far along, the M&E team instituted a periodic portfolio risk review. Since 2013, MCC M&E has conducted such reviews almost annually, using a standardized questionnaire. Questionnaires are completed by the staff managing each evaluation and then reviewed by M&E management. Through this process, the M&E team identifies:

- **Necessary course corrections for indi-vidual evaluations or a sector evaluation portfolio,** such as improved alignment of measurement plans and ex-ante expectations for the timing and magnitude of results, better coordinated indicator definitions, or measurement approaches, within a sector, and canceling and re-competing evaluations led by poor performing contractors.
- **Learning related to enhanced organiza-tional effectiveness,** such as how to improve coordination across M&E, sector operations and economic analysis for reducing evalua-tion risks.
- **Necessary updates to M&E guidance,** such as Evaluation Management Guidance revisions.

The first risk reviews addressed early evaluation weaknesses and brought them in line with M&E Policy and the Guidance. By 2019, the risk review

became more forward-looking, anticipating eval-uation challenges as early as the program design stage. The tool now assesses key elements of program design and implementation that facilitate quality evaluation. It is organized into five risk categories:

⚠ **Country** – factors outside MCC, MCA, or evaluator control

⚠ **Program design** – clarity and complete-ness of project design as the foundation for evaluation

⚠ **Program management** – adherence of program implementation to design

⚠ **Evaluation design** – alignment between evaluation design and project design/implementation

⚠ **Evaluation management** – adherence of evaluation implementation to design and guidance

This new tool also documents the use of evaluation evidence in program design or economic analysis. Around the same time, risk reviews shifted from being organized around country programs (i.e. all evaluations for a specific program are reviewed together), to being organized by sector to maximize coordination and learning within a sector.

Evaluation Leads

Beginning in 2013, M&E management designated Evaluation Leads for each common sector of evaluations. They oversee the quality of evaluation

work, enforce guidance, serve as hubs of sector knowledge to coordinate and improve mea-surement within a sector, and liaise with MCC's

complementary sector practice groups to promote sector-specific learning.[18] The growth of the M&E team's evaluation portfolio, the challenges in early evaluations that were frequently identified very late due to lack of standard oversight, and the need for a mechanism through which to enforce team guidance warranted the creation of this role. Evaluation Leads manage the quality of sector evaluation portfolios by mitigating risks and advancing best practices.

2.3 ADVANCING EVIDENCE USE (2017-PRESENT)

Evaluations show the extent to which a project achieved what it intended to and explain why or why not. This is the first step in **learning**. The second is to apply this knowledge to project decision-making. By 2017, the M&E team's evaluation practice had settled into a more reliable evidence production system with structures preventing many of the pitfalls experienced early on. This foundation, coupled with the fact that MCC had already published 68 interim and final evaluation reports, allowed the M&E team to reflect on the impact of the evidence and dissemination efforts thus far. This spurred a shift in focus from evidence generation to promoting evidence use.

This phase represented significant advancements in how the M&E team operationalized the principles of **learning** and **transparency** by further emphasizing access and use of evidence by internal and external stakeholders. This shift required increased attention to the end of the evaluation lifecycle, including the structure of results reports, the approach to dissemination, and the uptake and application of evidence at MCC and in partner countries. M&E staff identified the distinct users of independent evaluations and tailored unique evaluation products to them. In assessing the evaluation lifecycle, M&E staff identified various opportunities to better facilitate the use of evaluation evidence to inform decision-making at MCC and in partner countries. M&E staff continue to innovate in pursuit of this goal.

Final Evaluation Reports

Given that MCC's evaluation practice emerged from the agency's belief in evidence to improve development impact, MCC evaluation reports are highly technical in content, putting almost equal emphasis on analytical methods versus results. In addition, the broad and complex nature of MCC projects necessitates a large amount of data collection and analysis to assess the full set of targeted results and can lead to long evaluation reports. M&E staff, working with evaluators, have experimented with ways to make these technical reports more accessible, particularly to project teams and relevant policymakers, including:

- Presenting findings in line with the project logic - from outputs through outcomes sequentially - to pinpoint strong or weak causal linkages. This speaks to audiences with a technical interest in project design in a particular programmatic area.

---

18 Common sectors: Agriculture and Irrigation; Capacity Building; Education, Energy, Finance, Investment, & Trade; Health; Land; Transport; and Water, Sanitation, and Hygiene

- Using visuals to present results in a more digestible manner for any reader, including logic diagrams that use colors or symbols to indicate where the logic held and where it did not.

- Tailoring sections of the report to certain audiences and their need for detail, for example writing the executive summary for practitioners working in the same sector who require some but not all the details specific to MCC's program or putting technical details in an annex for audiences interested in the mechanics of the evaluation.

Evaluation Briefs[19]

In 2012, the M&E team began authoring the Summary of Findings to recap the evaluation's results and document lessons for MCC's practice. By 2017, M&E staff recognized these text-heavy summaries were still not optimally facilitating use by decision-makers and developed a new dissemination product, the *Evaluation Brief*. Launched in 2019, Evaluation Briefs are co-authored with the evaluator to maintain the integrity of the report's findings. They have strict length and formatting requirements that result in a more visual, readable, and predictable distillation of results, lessons, and methods. They are designed to serve various readers' needs. The full brief offers a comprehensive overview of the findings of the evaluation for MCC staff or the public. The standalone first page summarizes key results for policymakers or managers. The brief's heading and sub-heading convey the focus of the project and the evaluation's main takeaway for managers requiring succinct talking points. They also direct readers to MCC's wealth of additional resources with links throughout,
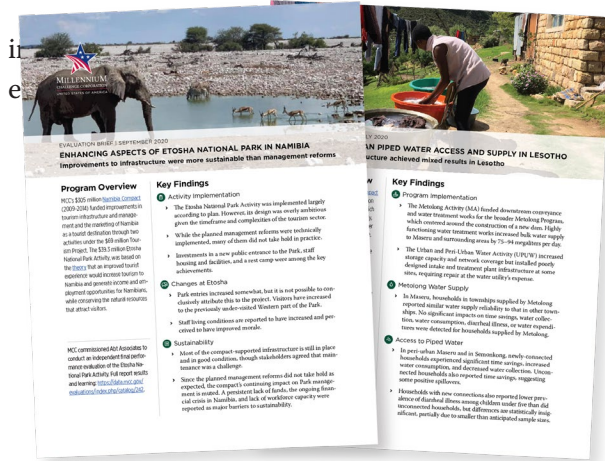
Evaluation Briefs are now widely used at dissemination presentations, meetings with high-level US and partner government stakeholders, and as references during new project design. The standardized distillation of results makes it easy to pull evaluation evidence for other materials, without requiring additional effort by M&E staff. Using Evaluation Brief information, the M&E team now sends periodic emails to the agency compiling evaluation results and lessons learned, called the Evaluation Dispatch.

MCC Learning

Beginning in 2012, as part of drafting the Summary of Findings, the M&E team introduced a standard approach to reflect on evaluation results, identify lessons for future MCC programming and evaluation, and document how they have been put into

practice at MCC, when applicable. Prior to this, there was no formal and routine means of documenting and publishing lessons from MCC's experiences, and the M&E team led with its core principles to embed **learning** in the evaluation lifecycle.

---

19   *Evaluation Briefs | Millennium Challenge Corporation (mcc.gov)*

M&E staff now convene a formal learning meeting at the finalization of every evaluation report to identify lessons for both future projects and future evaluations. The meeting includes the Evaluation Management Committee plus focal points to promote agency learning, such as the practice lead managing the sector's program portfolio. The

output of the meeting is the MCC Learning document, a description of each lesson and its motivation in enough detail for future staff to understand. These lessons are also aggregated in a searchable public database (see Sector Results and Learning Pages below).

Evaluation Dissemination Workshops

Since 2013, the M&E team has required dissemination workshops at both MCC headquarters and in partner countries at the conclusion of each interim and final report. This ensures all stakeholders have an opportunity to engage with the evidence produced by the evaluation. MCC funding for evaluator travel to partner countries for dissemination reflects the priority placed on broad engagement and use, as well as mutual accountability for results.

For many years, these events were academic-style presentations of results. However, starting in 2018, the M&E team shifted to a workshop format for

the MCC presentation that equally emphasizes the evaluation findings and the MCC Learning. This format calls for synthesized responses to the evaluation questions and presentation of results in line with the posited theory of change. It then brings in Committee members to present the MCC learning based on the evaluation and lead a discussion about what the evaluation means for the agency's practice going forward. This shift has resulted in more engagement from broader MCC staff at the workshops. The M&E team intends to make this shift in upcoming in-country presentations as well.

Sector Results and Learning Pages[20]

The M&E team launched the public *Sector Results and Learning Pages* in 2021 to further encourage the use of results and learning produced through MCC's dedicated investment in M&E. The pages are organized around MCC's six main sectors:
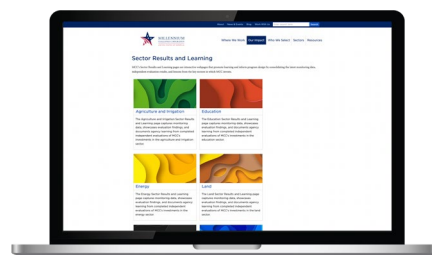

* Water, Sanitation, and Hygiene (WASH)
* Transportation
* Agriculture and Irrigation
* Education
* Land
* Energy

The pages map programs and common interventions in the sector, provide aggregated data on key performance indicators in each sector, and offer a curated set of evaluation results along with a

database of sector program lessons identified through the MCC Learning process. M&E staff tag these lessons with key topics related to the M&E team's risk assessments, like program design, program management, evaluation design, and evaluation management. This allows anyone accessing MCC's website to filter evaluation-driven lessons, based on their needs.

---

20   *https://www.mcc.gov/our-impact/sector-results*

Evidence Platform

Starting in 2013, the M&E team began sharing the reports, data, and other documentation produced by its independent evaluation portfolio on the Evaluation Catalog. It used open-source software to support MCC M&E's basic requirements for **transparency** regarding MCC program results and MCC's commitment to open data, however it offered limited MCC-specific search functions. [21]

In 2022, MCC launched a new *Evidence Platform* to respond to the usability requirements of various stakeholders – including country partners, students, researchers, congressional staffers, and the US taxpayer. The new platform is more user-friendly, with search, filtering, and sharing functions, among other improvements. It allows users to search the evidence by various types of knowledge products, including evaluation reports, public and restricted use data sets, data collection materials, and lessons.
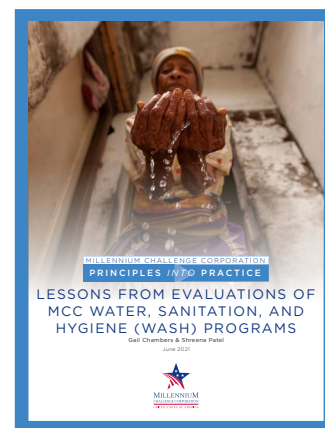
Star Reports

In 2019, MCC launched a new country program-level document, the Star Report, which describes the entire program lifecycle, from partner country selection to final project evaluation. While these are not strictly evaluation products and provide a broader narrative about the entire program that is authored by the program team, they include the high-level evaluation results and lessons drawn from the Evaluation Briefs and MCC Learning. The Star Reports serve as an important entry point to evaluation results for audiences that are less likely to seek out evidence through the more M&E-oriented avenues described above. This audience may include members of Congress or MCC budget appropriators who are most interested in a program-wide summary. The inclusion of independent evaluation results and lessons in these reports helps to ground the results narrative for each program in independent, publicly available evidence.

Principles into Practice Series

MCC produces sector-specific evidence and learning synthesis papers in its *Principles into Practice* series and other publications. Reports are undertaken when a critical mass of evaluations related to a specific sector or type of intervention has been completed. These papers are co-authored by M&E staff and the relevant sector experts. They reflect on the agency's experiences implementing and evaluating MCC programs and the evaluation

MILLENNIUM CHALLENGE CORPORATION
PRINCIPLES *INTO* PRACTICE

LESSONS FROM EVALUATIONS OF
MCC WATER, SANITATION, AND
HYGIENE (WASH) PROGRAMS
Gail Chambers & Shreena Patel
June 2021

---

21   Evaluation Catalog open-source software: *https://nada.ihsn.org/*

results, and draw broader lessons to improve future sector programming and M&E. Since 2012, four of these reports have been published: _WASH_ (2021), _Technical and Vocational Education and Training_ (2020), _Roads_ (2017), and _Agriculture_ (2012), along with an Issue Brief on _General Education_ (2023). Each Principles into Practice paper reflects MCC's commitment to **learning** from the measurement of results. They document how specific evaluation results or implementation lessons have changed MCC practice. For example:

- The agriculture paper influenced MCC's evaluation practice, and how MCC approaches coordination of hard and soft interventions in agriculture and irrigation projects.

- The roads lessons resulted in a restructuring of all existing roads evaluations and influenced investments in Nepal and Cote D'Ivoire to prioritize road maintenance over rehabilitation.

- MCC's WASH projects in Mongolia and Timor Leste reflect prior WASH evaluation lessons about the careful framing of project objectives and coordinating with water utilities on data collection.

Additional examples of how evaluation evidence has been used to improve project design at MCC, beyond what is captured in the Principles into Practice papers include:

Evidence in Action

Tanzania Compact
Transmission & Distribution (2017)

Evaluation provided evidence on the slower than expected rate of customer connections to the Compact's new electricity distribution lines and on the deterrents to connecting, such as high connection fees and complicated application procedures. In response, MCC rethought its approach to designing infrastructure programs **with more consideration for addressing the barriers to use of new infrastructure** through incentives or behavior change interventions. It also emphasized the need for more realistic assumptions about customer demand and uptake to inform project design.

Indonesia Compact
Green Prosperity (2018)

Findings from the implementation study of the Green Prosperity grant-making facility demonstrated the need for **problem-driven design of such facilities, a focused scope, and upfront market analysis and planning**. These points were reflected in MCC's guidance on designing and implementing leveraged grant facilities.

Burkina Faso Compact
Diversified Agriculture and Water Management (2021)

Evaluation demonstrated that long-term operations and maintenance of irrigated perimeters depend on payment enforcement and savings mechanisms and that the standard water user association model may not be sufficient. This lesson was applied to the irrigation project in the second Lesotho Compact, where the **project design incorporates governance boards, as a complement to WUAs**, to support sustainable operations and maintenance of irrigation infrastructure.

3. HIGH-QUALITY EVALUATION: REFINING THE DEFINITION

Throughout the evolution of independent evaluation at MCC, as the M&E team grappled with defining and applying core principles, setting and implementing evaluation management strategy and systems, and promoting the use of the evidence produced, the core mission of M&E staff has been to produce as high-quality an evaluation as possible. What constitutes high-quality was articulated piecemeal, leading to different points of focus. For example, an early MCC evaluation may only have been considered high-quality if it was an impact evaluation. As the Guidance was put into place, independence became a big focal point of quality. As the TREDD Guidelines were put into place, transparency took on new meaning in quality evaluation and the ethics of the evaluation garnered more attention. Finally, in the evidence promotion phase, the concepts of focus, accessibility, and use became paramount.

These various facets of good evaluation were reflected in different pieces of guidance and practice, and then in the evaluations themselves but had not been articulated holistically. Just as projects need to articulate a vision of success to design what is needed to get there, MCC M&E needed a clear, comprehensive definition of high-quality evaluation. This section of the paper presents such a

definition.[22] It is intended to serve as the model that each evaluation strives to achieve. It makes explicit what M&E staff and evaluators need to carefully consider and do.[23]

MCC's independent evaluations hold MCC and country partners mutually **accountable** for achieving targeted results. They contribute directly to **learning** the reasons targeted results were or were not achieved. The products of these analyses are posted publicly to support **transparency**. But what characteristics does an evaluation need to meet these core principles? How does MCC define a quality evaluation?

A quality evaluation:

- ✓ answers two summative evaluation questions regarding project effectiveness.

- ✓ is conducted by parties independent of project design and implementation.

- ✓ uses rigorous and ethical methods to produce unbiased analysis that is meaningful for decision-makers.

Drawing on nearly two decades of MCC's experience in evaluation, this section details MCC M&E's current definition of a high-quality evaluation. The theme of this paper, and an underlying force in MCC M&E, is constant self-reflection, learning, and improvement. The following definition represents current thinking but will be refined over time. Each new evaluation strives to improve upon previous ones, so there may not yet be an evaluation that perfectly checks all the boxes in the definition. This definition serves as a guide or a goal to aim for. Continuous progress over time, supported by adaptations to the systems described above, will bring the agency's evaluations closer and closer to this goal.

---

22   This definition has been incorporated into the revised M&E Policy, expected to be approved in 2022.

23   While this definition evolved out of MCC's own reflection, it aligns with USAID and OMB guidance on quality evaluations.

MCC's Definition of a Quality Independent Evaluation

A quality evaluation answers two **summative** evaluation questions regarding project effectiveness:

1. **To what extent was the project implemented according to plan (in terms of quantity and quality of outputs)?**

2. **Did the project achieve its stated objective in the timeframe and magnitude expected, as documented in the current M&E Plan? Why or why not?**

It is conducted by parties independent of project design and implementation. It uses **rigorous** and **ethical** methods to produce **unbiased** analysis that is **meaningful** for decision-makers. Specifically:

Evaluation design must be based explicitly on a project logic that reflects the project design, as presented by MCC and partner country stakeholders in the program agreement. It should be specified before the evaluation is conducted. It must measure indicators that adequately capture each result. Indicators to be analyzed in the evaluation are included in the M&E Plan.

Evaluation teams must be comprised of qualified researchers with expertise in evaluation methods as well as sector operations. They must be completely independent from the project. These third-party evaluators are hired by MCC.

Evaluation methodologies must reflect program implementation rules and local context to produce the most rigorous analysis of results that is feasible and cost-effective.

Evaluation samples should reflect the target population and have sufficient statistical power to measure the expected effect of the project on the objective.

Data collection methods – whether through quantitative and/or qualitative methods, primary and/or secondary sources – must follow sector best practice and ethical standards, include sufficient quality oversight, adhere to local data protection laws, consider the time commitment required of survey respondents, and consider the health and safety of enumerators and respondents.

Final data collection timing should align with the selected exposure period required to achieve the project objective.

Data analysis should be clearly documented such that readers can follow how results were derived and understand how they should be interpreted.

Evaluation results should be clear, concise, and specific; speak both to evaluators and policymakers; and offer clear recommendations for how to apply the results to practice.

Evaluation work must be subject to detailed review by MCC and partner country stakeholders, including project implementers, to ensure the accuracy of measurement and reporting while maintaining the evaluator's independence.

Evaluation resources must be used prudently, and the products should be shared transparently to contribute to the broader development and research community.

3.1 ANSWER CORE SUMMATIVE QUESTIONS

The M&E team's guiding principle of accountability and its commitment to independence require summative evaluation: a data-driven means to assess a project's effectiveness at achieving targeted results after its conclusion. MCC evaluations are designed to focus only on the following summative questions:

> **To what extent was the project implemented according to plan (in terms of quantity and quality of outputs)?**

Project implementation frequently deviates from initial intervention designs. Evaluations must first validate whether the intended treatment was provided as planned. They should assess the process and progress of implementation only in a level of detail necessary to inform the analysis of expected outcomes in Question 2. There is no expectation that the evaluator will audit implementation for compliance. Its purpose is to ensure a shared and accurate understanding between MCC and the evaluator of the reality of the intervention(s) being evaluated.

> **Did the project achieve its stated objective in the timeframe and magnitude expected, as documented in the current M&E Plan? Why or why not?**

This is the core of a summative evaluation and is key to mutual **accountability** for MCC and its partners. It focuses on the project's objective statement, which defines project success. It is the central question of the evaluation. To achieve **learning**, this question also asks why the objective was achieved or why it was not, requiring the evaluator to analyze intermediate results in the project logic linking MCC's investments to the objective.

Final evaluation data collection must be designed to detect the magnitude of change to the objective-level indicator(s) in the timeframe that was expected by the project team ex-ante. This target value reflects the project team's definition of project success and sets the sampling and timing of final data collection. Both the magnitude and timing are drawn from ex-ante economic analysis and/or project design documents on which the investment decision was based.

What about other questions?

Given MCC's appetite for continual learning, there will always be a demand for information beyond the evaluation questions above. For example, project teams may be curious to learn how different sub-populations experienced project benefits. Or the team may want to know whether additional benefits not expressly targeted by the project were produced, or whether there were externalities.

The M&E team's experience managing evaluations with long lists of questions led to a realization that this produced difficult-to-interpret reports that did not effectively support **accountability** and **learning**. Such lists of questions also hold different projects to different standards depending on the questions asked. As such, the M&E team now focuses independent evaluations

on only the two core questions above. The hypothesis is that focusing resources on answering these questions very well, instead of spreading effort too thin across many questions, will result in higher quality, actionable evaluation reports. This also makes more judicious use of M&E resources.

However, independent evaluations are not the only means by which MCC generates **learning**. Mechanisms such as the competitive research grant program included in the first Mongolia Compact and the coordinated research program in the second Morocco Compact can meet this demand. It is important, for the sake of accountability, learning and evaluation quality, to keep MCC's independent evaluations focused on their core questions.

What about formative work as part of independent evaluation?

There is often a desire to test specific project elements for real-time learning to inform a final project design and/or scale-up decisions. This calls for formative evaluation, a type of evaluation that assesses whether an intervention is feasible, appropriate, and acceptable before it is fully implemented. This type of evaluation can be very useful to improve project design and potential impact. However, it must be kept separate from independent evaluation in line with the OECD DAC principles for evaluation referenced in Section 2.1. Formative evaluation puts the

evaluation team in a project design role as their recommendations will steer the direction of implementation. The evaluation team therefore has a vested interest in the project succeeding. While they may not be biased, the risk remains that they will appear so. MCC M&E has not undertaken any formative evaluations to-date.

Formative evaluation was attempted once at MCC on the Kosovo Threshold Program. It was built into project design but had to be canceled due to delays in project implementation.

3.2 APPLY RIGOROUS METHODS

With the advent of the "randomista" movement to evaluate programs via randomized controlled trials, rigorous became synonymous with impact evaluation methodologies. However, acknowledging that impact evaluation may not be feasible or fit to purpose in all cases, the M&E team believes the analytical method is only one element of rigor and that others are also critical for producing credible evaluation evidence. The team considers the most rigorous evaluations to have several additional characteristics:

Based on a fully designed project that…

To conduct a rigorous, quality evaluation, the project itself would ideally be sufficiently well designed and documented. This is not always possible, and quality evaluation is certainly possible without it, but it increases the value of the evaluation.

Project design provides the basis for **accountability** – an evaluation cannot clearly determine if a project did what it aimed to do if the project never articulates its aims precisely. A fully designed project that is ready to support a high-quality evaluation:

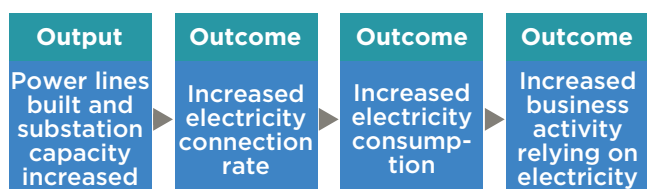### …articulates a clear objective ex-ante and defines it measurably

Each MCC project develops an objective statement that appears in the program agreement. This statement serves as the definition of success for the project, as determined by the project designers. The objective statement becomes measurable through precisely defined indicator(s) that articulate exactly what the project designers plan to achieve. M&E staff do not define these indicators on their own but work iteratively with the project team on the definitions to ensure they capture their vision of the objective precisely and measurably. For the sake of accountability, the objective must be articulated before a project (and certainly an evaluation!) begins. The objective should always be supported by a baseline value, demonstrating the project team appreciates the existing situation and that room for improvement exists.

…presents a pre-determined, evidence-based project logic

The project logic is the cause-and-effect theory explaining how the project designers believe the expenditure of project funds leads to outputs (the goods or services bought), intermediate outcomes (what happens because of the outputs) and the achievement of the project objective. It is the theory underpinning the project design, justifying the expenditure of funds. The theory must then be set out on paper before the evaluation can begin, otherwise the evaluation will not know what to measure and may waste resources measuring the wrong thing. It should demonstrate, in measurable terms, the results the project intends to deliver.

Output	Outcome	Outcome	Outcome
Power lines built and substation capacity increased	Increased electricity connection rate	Increased electricity consumption	Increased business activity relying on electricity

What if the project logic changes?

This is typically not a problem. Sometimes new information arises during implementation that requires a change to the project logic, such as adjustments to project activities or intermediate results. Sometimes this type of change is an explicit intention upfront, such as in the case of adaptive projects. Sometimes projects are approved with very simple logics, specifying little more than an objective and notional activities, and the logic must be fully elaborated during program implementation, such as projects that award grants. In all MCC evaluations, the evaluation design is guided by the project objective, which is specified upfront and will not change. It forms the core of the evaluation's measurement plans.

Evaluations of projects with changing logics, such as some policy and institutional reform projects, are no different than other types of projects. The only caveat is that if intermediate results change at the last minute, making ex-ante baseline data collection impossible, a retrospective baseline may be necessary. This could slightly limit the ability of the evaluation to speak to why and objective was or was not achieved.

...has a complete set of indicators to track the logic

A fully designed project defines each output and outcome in the project logic measurably. Each one must have a precisely defined indicator with clear targets. This information comes from project design documents, and/or the economic analysis. Like the project logic, these must be set out before the evaluation can begin so the right thing is measured. It is important, for the sake of both **accountability** and accurately capturing the project designers' aim, that these details are not developed by the evaluation team.

Applies a realistic evaluation design

Using the project logic and the indicators that measure it, evaluators design the most rigorous evaluation possible. In doing so, they assess the evaluation questions, rules of project implementation, and feasibility of various design options. Generally, this will call for a prospectively designed evaluation, or one that sets out the analysis approach before the project begins implementation.

While impact evaluations can more credibly attribute measured results to the project, such methodologies are not feasible for many MCC projects. Sometimes an appropriate counterfactual cannot be identified. This is due to the nature of MCC projects. MCC often funds large-scale investments that contain multiple activities aimed at producing change at a system-wide level. For example, projects may involve a national electric grid, or national road system, where entire networks are treated simultaneously. Additionally, the rules of implementation required to facilitate an impact evaluation or the cost and complexity of an impact evaluation relative to the expected benefits of the evaluation are factors that may also limit the appropriateness of this evaluation type.

When a performance evaluation is warranted, the M&E team requires it to be done as rigorously as possible. By triangulating data sources and carefully tracking the theory of change with data, performance evaluations may be able to build a credible case for or against contribution of the project to the measured change in the objective. This is especially the case when investments are large and timebound and the objective is more proximate to project activities.

Fit for purpose

While ex-ante baseline data collection generally facilitates more rigorous evaluation, this may not be realistic in all cases. Certain types of projects may not settle on a clear project logic until midway through implementation. This leaves evaluators with a limited window to design and conduct a baseline study. In such cases, a retrospective baseline designed after project implementation begins might be the most rigorous baseline feasible.

Uses evaluation methods and materials that are held to industry standards

In the evaluation design phase, evaluators must describe their analytical methods, clearly articulating the identification strategy for estimating project results, and presenting their analytical models. In the data collection phase, evaluators are expected to draw from international measurement standards or common survey instruments, such as UNICEF's Joint Monitoring Program for WASH or the Sustainable Development Goal indicators. In the reporting stage, evaluators must present both data tables reporting statistics, including p-values, and their interpretation of the data. Any claims must be clearly supported by the data collected. Results from impact versus performance evaluations must be caveated accordingly.

Assessing Results across Institutions, Enterprises, and Households

MCC's evaluations measure the results targeted by the project, which, given the diversity of MCC projects, means who or what the data will be collected on varies extensively. Depending on the logic of the project, MCC evaluations may need to be designed to assess policy reforms, organizational change, business productivity, and/or individual and household behavior change. As such, they may obtain data from businesses, utility staff, households, students, administrative records, roads, pipes, etc. One MCC evaluation will typically reflect many different data sources and evaluators must have the expertise to analyze various domains of results.

Has sufficient quality oversight

M&E staff provide technical feedback on all deliverables of an evaluation. This is critical to ensuring the rigor of evaluations is in line with MCC's standards. For example, in addition to reports and instruments, evaluators must provide documentation of pre-testing, translation quality checks, interviewer training, piloting, data entry, and data quality checks for review. This review also ensures these important steps are not overlooked. Evaluation Management Committee and partner country review strengthens this oversight, ensuring that the methodology and survey questions are appropriate for the local context and in line with project implementation.

Applies design and analysis specific to the sector and the local context

In validating the full theory of change of MCC projects, evaluators measure not only household and business level results, but also results related to performance of infrastructure, systems, and institutions. These measures, such as electricity outages, system water pressure, or road roughness, require technical sector expertise. Therefore, in addition to requiring expertise in evaluation, evaluation teams must also include sector experts, such as engineers when the project involves infrastructure, to ensure that the evaluation reflects an accurate understanding of the relevant measures and dynamics underpinning the project logic.

Similarly, local expertise is vital to a successful evaluation. Each MCC evaluation benefits from local stakeholder engagement on their design, implementation, and dissemination, from MCA, partner government, and local implementer input. Preliminary evaluation plans are part of the program agreement between MCC and the partner government and receive high-level buy-in. In addition, evaluators almost exclusively partner with local survey firms to ensure evaluation data collection asks the proper questions and is implemented appropriately for the local context.

Building data capacity for business needs

MCC M&E and sector experts work to ensure relevant government authorities get the support they need in collecting data relevant to project results. Such entities, for example a utility, are best placed to measure systemwide outcomes. They typically have a business need for the data as well. This requires collaboration across MCC and integrating the support into the project.

Appropriately balances rigor with cost

Some aspects of evaluation design might be proposed with the argument that they will bring more rigor. However, this proposal should be looked at carefully. Evaluation rigor can be diminished when evaluation resources are stretched too thin or when a well-resourced evaluation tries to respond to too many varied demands. For example, an evaluation may propose:

- More rounds of data collection than are truly necessary

- A more sophisticated data collection technology than what would get a satisfactory estimate
- Triangulating quantitative and qualitative data when only one is enough to understand the results.

M&E staff protect the rigor of an evaluation by carefully managing its design and the associated cost. As described in Section 2.2.2, the Committee is responsible for weighing expected **accountability** and **learning** benefits against the expected costs. The M&E team requires evaluators to detail and

justify all elements of the evaluation design (e.g., methodology, sample, sample size, rounds of data collection and timing, data collection methods and detail, detailed costs, etc.) in the Evaluation Design Report, alongside its costs. This provides the Committee with sufficient information to assess the rigor of the design and determine whether all elements are truly necessary. M&E staff oversight ensures that evaluation resources are put to their best use and that the evaluation remains focused on rigorously answering core questions.

3.3 CONSIDER ETHICAL RESPONSIBILITIES

Since 2013, MCC has called attention to the need for MCC-commissioned evaluations to integrate research ethics - respect for persons, beneficence, and justice – into evaluation design, implementation, and dissemination.[24] Additionally, MCC aligns with the Federal Data Strategy and is committed to open data and **transparency**, while

recognizing the need to balance this commitment with proper, ethical data management when evaluations rely on personally identifiable information (PII) and/or sensitive data.[25] For ethical data management, MCC aims to ensure both staff and evaluators:

Adhere to promises of confidentiality to reduce risk of direct harm to data providers

If unauthorized users obtain PII or sensitive information that is linkable to the data provider, there is a risk that this disclosure could be used to harm and/or exploit the data provider.

Adhere to promises of confidentiality to reduce risk of reputational harm to data handlers, country partners, and MCC

MCC staff and contractors must carefully consider and implement best practices to mitigate against unauthorized disclosure of PII and sensitive data of its respondents. Experience has shown this is not easy, as there is wide variation in ethical oversight.

There is no standard oversight body for MCC – or the broader international development community – which governs the ethical review of social science research.

---

24 For example, the M&E team learned that contractor staff were not necessarily required by their own organizations to be trained on the protection of human subjects and so this had to be a required qualification stated in the MCC evaluation contract.

25 Federal Data Strategy available here: *https://strategy.data.gov/principles/*

Per TREDD, the M&E team requires evaluators to go through Institutional Review Board (IRB) review. The assumption was that this process would ensure a comprehensive review of the ethics of each evaluation. However, this is not the reality. Many MCC evaluations are considered exempt from IRB review by the IRB itself. [26] Even when MCC insists on IRB review, there are many topics for which little evidence-based guidance is given by IRBs, such as interview duration limits, respondent compensation, appropriate content for informed consent, and determination of respondent comprehension of informed consent process. Additionally, about 25% of MCC evaluations are conducted in countries with no local IRB, so an IRB without local knowledge must be used. Finally, evaluation firms, data collection firms, and other key country partner

staff often require MCC support because they do not have resources for or access to standard training on protection of human subjects and principles of research ethics. With this in mind, MCC engages with legal counsel up front to understand local data protection laws and ensure that all evaluation work is compliant. Evaluation contracts stipulate the duration of surveys must be no longer than 2 hours, to be respectful of respondents' time, and evaluations consider respondent compensation in line with the local context. MCC aims to continue working with the international development community to consider more centralized, common frameworks for careful consideration of human subjects, particularly the health and safety of respondents and enumerators, and responsible, ethical data management.

3.4 ENSURE UNBIASED EVALUATORS

A challenge for M&E at MCC lies in how to ensure that any measurement, analysis, and interpretation of results is unbiased. It is important to ensure that incentives to exaggerate or overreport positive results or remove or underreport negative and/or null results been sufficiently reduced or mitigated. MCC seeks unbiased analysis of its results by commissioning independent evaluations by third party evaluators who are not involved in program design and have a shared interest in maintaining reputations for unbiased analysis.

Transparency is a key principle, as well as a tool for achieving an unbiased analysis. MCC and partner country stakeholders review all evaluation materials, starting with the evaluation design report and ending with the analysis report and data files, as a mechanism of holding evaluators accountable for producing unbiased results.

Reflecting the importance of evaluation at MCC, participation in these reviews is typically high. The report review and feedback process are transparently documented in the Guidance. All evaluation reports are published with reviewer comments and evaluator responses included as an annex. This holds reviewers accountable for what they write, as comments will be made public. Evaluators respond to each comment but are free to decide whether to make a change to the report. The existence of the Committees at MCC or MCA from the start of the evaluation ensures that evaluators receive comments on every report, either through written review or during presentations. Posting this dialog publicly demonstrates unbiased independence. De-identified data sets and analysis codes are also posted publicly to allow for replication of results and hold evaluators accountable for the quality of their analysis and reporting.

---

26   Under the Federal Policy for the Protection of Human Subjects, research that is designed to study, evaluate, improve, or otherwise examine public benefit or public service programs and is conducted by a federal department or agency, is exempted. IRBs have concluded that MCC-funded independent evaluations fall into this category.

Who are MCC's evaluators?

Since 2006, MCC has procured and awarded contracts to research firms and academic institutions specializing in evaluation of international projects. Early into the start of evaluation procurements, MCC identified pools of qualified firms and negotiated base agreements with them that allowed for a simplified contracting procedure for individual evaluations. Early on, roughly 20 contracted evaluations had to be canceled because of failed methodologies, poor performance, or a poorly defined scope. By clarifying MCC's evaluation needs for the market and emphasizing the independence and rigor of the evaluations, the Guidance and standard Scope of Work has helped MCC to attract high-quality firms to partner with. MCC has continued to work with many of the same research firms since the beginning. MCC has also found that working with academic institutions is better in partnership with one of these research firms.

3.5 PRODUCE MEANINGFUL WORK

Ensuring an evaluation is meaningful builds on and goes beyond the technical details discussed above. A meaningful evaluation is one that is used to inform future decisions at MCC, in partner countries, and maybe even across other development agencies. To ensure this happens, an evaluation must:

- Be relevant to decision-makers
- Be respected and trusted so decision-makers will view the results as credible and act on them
- Make results available in as timely a manner as possible, enabling the integration of learning before a similar intervention is implemented
- Be easily accessible to decision-makers, in terms of the method of delivery, structure, and content

MCC M&E's guidance and processes set a foundation for this, but it requires inclusive M&E staff leadership of the evaluation to achieve one that is truly meaningful.

While other aspects of the project logic related to rigor have already been discussed, for the evaluation to be meaningful, it is also important that the logic accurately capture the project designers' theory of the investment. Ensuring this happens, and the logic is updated as designs change, is critical for producing a relevant evaluation. The Committee provides a mechanism for ensuring communication between the evaluator and the project implementation team stays open and transparent. M&E staff must actively manage this relationship.

Stakeholders will not base decisions on the results of an evaluation if they do not buy into the evaluation design at its earliest stages, as is facilitated through the Committee reviews. Any concerns about the design can be brought up early on when there is time for change. This ensures that the results of the evaluation, whether they are positive

or negative, will be accepted and stakeholders can confidently act on them.

Meaningfulness also requires timely reporting on results. Given the focus on the project objective, which usually manifests a year or two after the MCC program has closed, it is generally not possible to present final evaluation results right as programs end, or very soon thereafter. The process cannot be sped up, but clear communication of the timeline and its rationale helps manage expectations. If there is demand for results sooner to inform a particular decision or milestone, it may be possible to conduct an interim evaluation of results earlier in the project logic, such as outputs or intermediate outcomes. [27]

Finally, to ensure analysis is used, its presentation must be accessible, both in structure and substance, to the intended users. MCC evaluations are clearly structured around a theory, which is tested and analyzed, making it easy to draw out practical lessons. The presentation of results is tailored to the needs of the intended audience. Many decision-makers will require a concise summary of the project and its results, with very clear, straightforward answers to the evaluation questions. Accessibility also means ensuring evaluation results, and the lessons learned from them, are readily available online in a user-friendly format to the broader development community outside of MCC, as provided by the Evaluation Briefs. The M&E team listens to feedback from users and continues to experiment with ways to enhance the clarity and ease of finding, searching, and using evaluation results.

While the approaches described throughout the paper facilitate meaningful evaluation, they do not guarantee it. Constant attention, communication, and collaboration amongst the members of the Committee is critical. Strong technical skills are a must for M&E staff, but crafting a meaningful evaluation goes beyond those hard skills, requiring M&E staff to be the true leaders of the Committee and the whole evaluation process.

Key Findings

Traffic Counts, Travel Time, and Costs

› The evaluation found lower than predicted traffic counts for the Highway CA-5 road segments but higher than predicted traffic counts for secondary roads.

› There were statistically significant reductions in travel times and costs.

Prices, Employment, and Income

› Monthly agriculture income rose, but monthly non-agriculture income fell.

› There was no statistically significant impact on overall household income or wage employment.

Road Condition and Maintenance

› The quality of the paved road upgrades and improvements was quite high and met the highest international standards.

› For all rural road upgrades, depending on traffic volume and road alignment, road conditions were expected to return to pre-improvement levels without further investment or maintenance.

› There were higher than expected road maintenance costs and final project costs, but these increased costs did not negatively affect cost-effectiveness.

---

27   Approximately 25% of completed or active MCC evaluations have produced an interim evaluation report.

4. MCC EVALUATION OVER THE NEXT DECADE

Going forward, MCC will continue to advance its mission to reduce poverty through economic growth. Its unique evidenced-based model, underpinned by the core principles of good governance, country ownership, a focus on results and transparency, provides a foundation for a high-quality evaluation practice to support achievement of this mission.

Building on this agency-wide foundation, MCC evaluations are guided by the principles of accountability, learning and transparency, with the ultimate goal of facilitating learning that improves the impact of future projects. This evaluation practice is also influenced by the various contexts within which MCC operates. MCC is a US government agency, which must adhere to federal regulations and administrative priorities. It is a partner to various country governments, beholden to their priorities and research and data privacy laws, and it is a development agency among bi-lateral and multi-lateral peers subscribing to internationally agreed best practice. As described in this paper, with each completed evaluation, the M&E team reflects on how well this goal was achieved, learning and changing practices as necessary to improve.

Considering the decade ahead, some key areas of potential evolution of evaluation at MCC emerge. These topics are likely relevant to other international development organizations. These topics provide an exciting opportunity for MCC to lead on how best to use evaluation, **accountability** and **learning** to improve international development practice.

WHAT IS THE "RIGHT" LEVEL OF ACCOUNTABILITY?

MCC projects consist of multiple, separable components working in concert toward a common aim. Should each project component be evaluated individually rather than speaking to project results as a whole? One may argue that, as a publicly financed entity, MCC has a responsibility to report on the specific effects of every component invested in. This information may also be useful to the teams implementing these components. On the other hand, should programs be evaluated as a whole, looking at the combined impact of its projects? This question may be of interest as well, especially to stakeholders overseeing MCC's work.

These views are valid but require a different allocation of M&E focus and resources. Currently, MCC M&E evaluates each project to align with the agency's legislation and the agency's centering of

country programs on distinct projects. This allows for a more efficient allocation of M&E resources to provide a clearer assessment of aid effectiveness, by project. However, if MCC seeks a more granular level of accountability, expanding evaluation scope may be justified; though this will result in more evaluations per program and requires additional resources. A broader level of accountability would also require a shift in evaluation scope and potentially more resources than a project-level approach, as program-level impacts may be very difficult to detect. It may also be worthwhile to consider broader developments in the community of practice around evidence generation and use, such as the effort around using evidence to identify the core components of program effectiveness. Aligning MCC's evidence-generation approach with such initiatives may also necessitate changes.

IS THE PROJECT OBJECTIVE THE RIGHT ANCHOR FOR EVALUATIONS?

As described in Section 2.2.1, "Refining the meaning of accountability in MCC evaluation", MCC M&E evaluates each project on achievement of its stated objective, because that objective codifies the definition of project success and is stated prominently in the first section of each program agreement. However, there may be additional aims that are not explicitly captured in the objective and associated project logic of achieving it, such as standalone results of sub-project components or results posited by the cost benefit analysis (CBA) that go beyond achievement of the objective. There is also the theory beyond the project objective, of how the project contributes to the goal of all MCC investments - poverty reduction through economic growth.

One may argue that these results are also important to measure for the sake of accountability, in addition to the project objective. Standalone results may represent agency priorities that influence all projects and for which the agency is accountable. Results in the CBA reflect the economic rate of return that informs the investment decision. MCC evaluations that focus on the objective may miss an opportunity to measure MCC's impact on poverty or to contribute to the literature on the economic impact of development interventions. Like the above question, expansion of evaluation scope with additional resources may be possible and may be more desirable considering MCC's current strategic priorities around climate change, blended finance, and gender and social inclusion.

IS A FOCUS ON ACCOUNTABILITY ADDING VALUE FOR MCC?

MCC M&E strives to add value to the MCC model, and so it is worth reflecting on the benefits of the accountability focus. Accountability is an excellent goal, and one that the agency prides itself on. It is important to learn the actual impact of our projects – sometimes such a strong project logic narrative exists that it is easy to forget that one does not yet know how that theory played out in real life. The evaluation provides that truth, or accountability, and will inevitably force the agency to grapple with negative results and learn.

One challenge is that for accountability to add value, project objectives need to be expressed precisely and realistically, as also required by MCC's statute. There is always a tension between setting project objectives that sound impressive versus those that the project team is confident they can achieve. There is also a tension between setting project objectives in general terms versus setting out precise targeted metrics. If the project team chooses a project objective they know they are unlikely to be able to achieve, or doesn't precisely define it , the accountability exercise loses value.

More broadly, it is worth considering what accountability means or should mean, comprehensively. Is measuring the success of a project through evaluation enough? What needs to happen after that measurement to claim full accountability exists? The MCC M&E model assumes a major value-add of accountability is that it results in learning. This brings us to our next question.

CAN MCC'S MODEL BE OPTIMIZED FOR LEARNING?

In recent years, MCC has strived to strengthen the feedback loop between evaluation results and project decision-making. In doing so, some interesting points have come to light, points that are relevant to any organization that completes evaluations and wants to use the results.

First, how does an agency remember all the lessons identified through the course of evaluation so that it can truly learn? As described above, MCC M&E has created a database with all evaluation lessons, tagged and easily searchable. Still, it is very possible a lesson might get overlooked during project development. There are hundreds, and the database is growing.

Even if relevant lessons are identified, they can be hard to implement. Hopefully the Evaluation Management Committee approach described above ensures that the appropriate staff members find the evaluations relevant and applicable. But is an evaluation lesson enough to justify to a stakeholder why an intervention cannot be pursued? It is a difficult conversation. Project development teams also face time pressure, making gathering the necessary evidence and carefully considering it hard. Are there ways MCC's model, particularly its project development processes, can be tweaked to facilitate learning from evaluations?

DOES MCC'S EVALUATION PRACTICE STRIKE THE RIGHT BALANCE BETWEEN ACCOUNTABILITY AND LEARNING?

MCC evaluations take a very specific approach, the historical evolution of which is described in this paper. However, MCC M&E can be flexible about the evaluation approach to be taken going forward and is open to changing course and shifting its principles if that is deemed to be of greater value to the agency. To learn about broader set of results, perhaps it would be more valuable to focus M&E resources on exploring all possible results rather than solely those targeted ex-ante by projects or those supported by existing evidence . It may also be more useful to produce evidence-based success stories to learn about what worked well at any level of the theory of change. This would require a different set of M&E principles, a different approach to designing evaluations, and possibly a different role for the evaluator.

Another topic of discussion at MCC is formative evaluation or other real-time work to inform the direction of project design. As described in in Section 3.1 ,"What about formative work as part of independent evaluation?", this approach is not compatible with independent evaluation, but is still possible to do at MCC. As it requires a detailed knowledge of the project and the ability to make and implement quick decisions based on the data, it is best managed by project teams themselves. But it is something MCC M&E staff and resources could support.

These topics are all substantial. They may not all be tackled in the next decade but should spark some interesting conversations. These conversations are relevant to the broader international development and evaluation community and are conversations that MCC is well-placed to lead. MCC's values, its evidence-based model, and its unwavering support for evaluation put the agency at the forefront of the evaluation field. Beyond the outcome of these questions, MCC M&E has evolved substantially over time and will continue to do so. It has been eight years since MCC last reflected publicly on its evaluation practice, and this paper demonstrates the considerable advancement of M&E at MCC since then.

Reducing Poverty Through Growth