

The Balanced Design, Monitoring, Evaluation, Research, and Learning (BalanceD-MERL) Maturity Matrix

December 2018



This publication was produced by the William Davidson Institute at the University of Michigan (WDI), Institute for Development Impact (I4DI), and World Vision under the BalanceD-MERL Program, Cooperative Agreement Number AID-OAA-15-00061, funded by the U.S. Agency for International Development (USAID). This work was made possible by the generous support of the American people through USAID. The contents are the responsibility of WDI, I4DI, and World Vision and do not necessarily reflect the views of USAID or the United States Government.

ABOUT BALANCED-MERL

The Balanced **D**esign, **M**onitoring, **E**valuation, **R**esearch, and **L**earning (BalanceD-MERL) consortium recommends deliberate balanced integration across all aspects of D-MERL to enable better use of monitoring, evaluation, research, and learning (**MERL**), leading to a team's ability to gather and incorporate findings into program design (**D**). This can create sustainable improvements in the program, and lasting positive impacts on programs' target audiences. Balance can be achieved by explicitly incorporating four key principles (see **Box 1**), underpinned by a focus on utilization, into D-MERL activities. Through the use of these principles, staff can iteratively learn from program implementation, practice good program management and evidence-based decision-making, and incorporate lessons into program design and implementation.

Box 1: Definitions of the four principles of BalanceD-MERL

- 1. **Relevant:** D-MERL is relevant when it is informed by development theory and is intentionally shaped by, and responds to, how local people, context, and strategy evolve over time.
- 2. **Right-sized:** D-MERL is right-sized when it is a match between resources (people, time, and money) and goals. D-MERL activities, processes, and methods are coordinated and efficient. D-MERL processes and data are non-duplicative and meet program management, leadership, and compliance needs.
- 3. **Responsible:** D-MERL is responsible when it goes beyond the "do no harm" principle to engage respectfully, ethically, and sensitively with the target audience(s) of the program and local partners. D-MERL is transparent and accountable.
- 4. **Trustworthy:** D-MERL is trustworthy when it is conducted according to standards of rigor appropriate to context, constraints, and/or intended use of the data. Trustworthy D-MERL is internally and externally valid, reliable, and objective, ensuring integrity of both the process and the results.

The BalanceD-MERL consortium under the U.S. Global Development Lab's Monitoring, Evaluation, Research, and Learning Innovations (MERLIN) program consists of World Vision (Prime), Innovations for Poverty Action, Institute for Development Impact, Search for Common Ground, and the William Davidson Institute at the University of Michigan. More information on the BalanceD-MERL approach, including case studies of two pilot projects leveraging this approach, can be found here. The Maturity Matrix described in this document has also been informed by the pilot projects.

U.S. Global Development Lab contact: Shannon Griswold

WHO CAN USE THE BALANCED-MERL MATURITY MATRIX?

The BalanceD-MERL Maturity Matrix can be used by program and Monitoring, Evaluation, Research, and Learning (MERL) staff who facilitate program design (D) and implementation and its performance management and/or evaluation. This can also include any persons who have responsibilities for steering the course of a program. The matrix can be used by staff involved in a single program or project with a single MERL plan, or it can be used with a complex portfolio of programs/projects that may be aligned under a broad MERL strategy.

WHAT IS THE BALANCED-MERL MATURITY MATRIX?

The matrix is a tool to help program and MERL staff integrate MERL with program design through the use of four principles (see **Box 1** above) in the service of good program management. Through these principles, activities, programs, and strategies are informed by:

- 1. Development theory and shaped by the local context
- 2. Stakeholder engagement is guided by culturally-sensitive and ethical practices
- 3. Resources are appropriately matched with program and strategic goals
- 4. Valid, reliable, and objective results are found and
- 5. Decision-making is evidence-driven and reflective of a broad range of stakeholders' voices.

The matrix also provides a framework to learn from, adapt, and improve both program design and its implementation and MERL design and its execution. All this enables *best-of-class* and *evidence-based* adaptive management throughout a program's lifecycle.

The rows of the Maturity Matrix are organized by the four principles, with multiple rows of content dedicated to the various aspects of each principle. The columns are organized by three levels of maturity for each principle (*unbalanced*, *emergent*, and *balanced*) which show the progression of D-MERL activities and processes along the spectrum. Each cell contains a description of D-MERL practice for a given aspect of a principle and at that particular level of maturity, followed by guiding questions to help move users to the *balanced* state. A program's D-MERL maturity may vary between and within the different principles. The variation can guide where a program's D-MERL practice can be improved.

A first step in using the Maturity Matrix is for users to self-assess the D-MERL maturity of their program across the principles. To do so, users may find it helpful to go through the Maturity Matrix row-by-row to identify what level of maturity their program is in for each row. Once a self-assessment is completed, the user can review the guiding questions in the cells that represent their current state to formulate a plan for how to move towards the *balanced* level of maturity. The guiding questions are meant to help users mature their D-MERL practice, where needed, bearing in mind that balancing may not be possible all at once. The guiding questions are not intended to be exhaustive nor are they meant to be diagnostic. The questions help users navigate between the levels of maturity by offering prompts and, in some cases, tools and resources to improve D-MERL practice. The user is still required to read the description above the set of guiding questions in each cell and think about where their program falls on the spectrum, and *why*. Once a user understands their program's level of maturity for a given principle, they can use the questions and conduct additional research on how to put the principles into practice, within their own context. Thus the Maturity Matrix allows users to:

- 1. Self-assess their program(s) maturity for each principle
- 2. Reflect on their program's current status and develop future goals on D-MERL integration and adaptive management
- 3. Benchmark their program against good MERL practices that are well integrated with program design and implementation

WHAT WILL ENABLE USE OF THE MATRIX?

These are two suggestive factors that can enable use of the matrix:

- 1. Leadership buy-in and readiness to use the matrix and the larger BalanceD-MERL approach including providing staff the necessary resources (especially time needed) to practice the approach, and
- 2. As applicable, flexible donor contracting and reporting processes to support the practice of adaptive program management.

HOW WAS THE MATRIX DEVELOPED?

The matrix was developed by the BalanceD-MERL consortium through a multistep process beginning with a literature review. The consortium explored what constitutes each principle as well as what aspects of each principle should fit within each level of maturity. Then the consortium refined the content based on internal learning and feedback from reviewers in the consortium's member organizations, as well as external monitoring and evaluation practitioners.

USEFUL DEFINITIONS

- Adaptive management is defined as an intentional approach to making decisions and adjustments to improve the
 program's activities and implementation in response to new information and changes in context that is captured through
 MERL data. Here decision-makers intentionally seek to incorporate learning into program delivery. This definition of
 adaptive management is based on the definition provided by USAID in ADS 201.6.
- **Data quality:** The five dimensions of data quality are the following: **valid data** clearly and adequately represents intended results; **reliable data** reflects stable and consistent data collection processes and analysis methods over time; **timely data** is available at a useful frequency and timely enough to influence leadership decision-making; **precise data** has a sufficient level of detail to permit leadership decision-making; and having **integrity with data** means that the data has passed through safeguards to minimize risk of transcription error or data manipulation (as <u>defined by USAID</u>).
- D-MERL is defined as the integration of MERL activities with program design in the service of good program management.
- **Indicator metadata** includes: indicator definition, unit of measure, source of data, reporting frequency, the method to collect it with limitations and actions to address the limitations, and how the indicator will be disaggregated.
- **Key stakeholders** are defined as (1) primary intended users, (2) key local partners such as local non-profits, civil society organizations, and government, (3) organizations implementing program activities, and (4) funders of a program. These multiple and varied individuals and groups may affect or be affected by the program.
- **MERL activities** are both the activities and the iterative process of developing activities for performance measurement of the program. These activities measure and monitor how effective a program is at a moment in time as well as how to make the program more impactful in the future. Key activities include developing and implementing the following: (1) theory of change including program assumptions, (2) risk matrix, (3) evaluation matrix including the learning questions, (4) key performance indicators (with relevant metadata), (5) research design including sampling strategy for the treatment and comparison group if applicable, and (6) data analysis plans and methods.
- Organization and support services refers to other units within an organization, such as finance and accounting, procurement, and human resources.
- Program design are both the activities and the iterative process of developing activities to achieve the objectives and goals of the program.
- **Primary intended users** are beneficiaries in a non-market-based program or customers, distributors, producers, and/or employees in a market-based program.

Lev	vel 1: Unbalanced application of the principle	Level 2: Emergent application of the principle	Level 3: Balanced application of the principle			
Relevant: D-MERL is relevant when it is informed by development theory and is intentionally shaped by, and responds to, how local people, context, and strategy evolve over time						
Relevance – Aspect 1	Key stakeholders are not identified. There are no efforts to engage any key stakeholders to account for the local context, throughout the life of the program. This information is not consulted in the design and implementation of D-MERL activities and processes.	Key stakeholders are identified but are consulted in an ad-hoc manner throughout the life of the program to account for the local context in the design and implementation of D-MERL activities and processes. The information received in the consultation is reviewed in an ad-hoc manner. The feedback and needs expressed by these stakeholders receive some but mostly token consideration.	Key stakeholders are consulted routinely throughout the life of the program in the design and implementation of D-MERL activities and processes to account for the local context. The feedback and needs of these key stakeholders are prioritized. The feedback is acted upon to improve D-MERL for increased impact. Their voice is included when making decisions on changes to D-MERL activities and processes.			
	Guiding questions to the next level of the matrix 1. Have you determined who are your key stakeholders? A stakeholder assessment can help.	Guiding questions to the next level of the matrix 2. Are you providing opportunities for key stakeholders to engage with and provide input to your D-MERL activities and processes at critical points? You could do this by working with individual point of contacts of your key stakeholders or setting up a community advisory board . 3. Are you taking the time to review feedback from key stakeholders?	4. Are you prioritizing feedback from key stakeholders in your D-MERL activities and processes? 5. Beginning from program design and continuing throughout the life of the program, are you routinizing opportunities for key stakeholders to provide input to your D-MERL activities and processes such as once a year or more frequently as required? 6. Are you routinizing the opportunities to review and incorporate feedback from stakeholders?			
Relevance – Aspect 2	The program staff and leadership are not using measurement data and insights from MERL activities and processes to adjust the program or future programs based on changes to context.	The program staff and leadership are reviewing MERL data and insights retrospectively to detect changes in internal and external context. However, they still do not adjust the program over time or in time and/or the design of new programs. MERL activities and processes are also not updated based on the MERL data and insights gathered. Consulting MERL data is not considered as part of the formal roles and responsibilities of any staff members.	The program staff are using MERL data and insights to detect changes in internal and external context in realtime and make any necessary adjustments to the program i.e., decision-making on program changes are driven by MERL data. New program designs as well as MERL activities and processes incorporate MERL data and insights when necessary. When using the MERL data for decision-making, program staff also recognize the limitations of the MERL data (if any). Consulting the MERL data and insights is routinely carried out throughout the life of the program. These are documented along with the changes to the program. The routine consultation of MERL data and insights are incorporated into formal definitions of roles and responsibilities for all staff.			
	Guiding questions to the next level of the matrix 1. Do you have a current understanding of your operating environment? A context analysis can help. 2. Do you have a current understanding of the needs of your target population? A needs assessment can help. 3. Periodically, do you take time to review your MERL data and insights to identify necessary adaptations to the program?	Guiding questions to the next level of the matrix 4. Are you analyzing and learning from MERL data in real-time? Do you routinely take time to adjust your program based on real-time data?	Guiding questions to achieve this level of the matrix 5. Do staff job descriptions make everyone responsible for the identification and use of MERL data and insights? 6. Have you identified the limitations of the data to ensure you do not overextend the use of the data?			

efforts because no literature and secondary data reviews are carried out to inform D-MERL design.

Guiding guestions to the next level of the matrix

- 1. Have you identified the financial resources available for D-MERL activities and processes?
- 2. Have you identified the staffing requirements based on technical skill and capacity needed to design and implement the D-MERL activities and processes?
- 3. Have you identified timelines available for D-MERL activities and processes?
- 4. Have you conducted a literature review of what data and insights are already known? Do your D-MERL activities and processes account for the known data?

Guiding questions to the next level of the matrix

MERL activities and processes.

5. Have you developed your D-MERL activities and processes based on the resources you have available (financial resources, staff capacities, timeline)? 6. Have you developed your D-MERL activities and processes such that they can measure the results in your theory of change as well as program objectives and impact goals?

5. Do you always apply MERL activities and processes as they are documented and are these coordinated

D-MERL activities and processes align to resources (people, time, money) and the theory of change. The D-MERL activities address the program objectives and literature review is conducted including speaking with experts (if possible) to eliminate duplicate efforts.

Guiding questions to achieve this level of the matrix

7. Have you identified the gaps in resources and determined steps to fill those gaps such that you can carry out the necessary D-MERL activities and processes?

Le	vel 1: Unbalanced application of the principle	Level 2: Emergent application of the principle	Level 3: Balanced application of the principle			
R	Responsible: D-MERL is responsible when it goes beyond the "Do No Harm" principle to engage respectfully, ethically, and sensitively with the primary intended users and local partners. D-MERL is transparent and accountable.					
le – Aspect 1	MERL activities and processes are not sensitive to key primary intended users' needs and to local social and cultural constructs. These MERL activities and processes add to key primary intended users existing burdens. None of the data collection tools have been pretested. Leadership, program, and MERL staff do not recognize ethical considerations and therefore, overlook these.	MERL activities and processes partially acknowledge the sensitivities of key primary intended users' needs and social and cultural constructs. The burden on primary intended users is recognized but not fully addressed. Data collection tools and processes have been poorly pretested. Leadership, program, and MERL staff have been given limited training in ethical considerations. The ethical considerations of all MERL activities and processes has not been assessed.	MERL activities and processes are sensitive to key primary intended users' needs, and social and cultural constructs. The burden on primary intended users is recognized and fully addressed. Data collection tools and processes have been pretested. Leadership, program, and MERL staff are given training on ethics and sensitivity and apply these lessons in all MERL activities and processes.			
Responsible	Guiding questions to the next level of the matrix 1. Have you determined how key norms (e.g. gender) in the communities you are working with will influence MERL activities and processes? 2. Have you reviewed your MERL activities and processes with relevant stakeholders to determine if any of these create burdens on key primary intended users and program staff?	Guiding questions to the next level of the matrix 3. Have you determined if an incentive is appropriate to compensate for the primary intended user's time? 4. Have you conducted high quality pretesting of data collection tools with a representative sample? 5. Have you conducted trainings on ethics of research with your leadership, program, and MERL staff as well as key implementing partners?	Guiding questions to achieve this level of the matrix 6. Are you monitoring and updating MERL activities and processes according to any ethical considerations that arise during program implementation and MERL data collection?			
- Aspect 2	There is no sharing of data, results, and/or limitations of the data either internally with key stakeholders or externally with the broader global development community. There is no formal or informal policy on internal and external sharing.	Data and results are shared both internally with key stakeholders and externally with the broader global development community, but any data and results that are perceived to be negative are redacted. There may be sharing of some limitations. Formal policies on sharing both internally or externally have been developed but are inconsistently implemented.	Data and results sharing is widespread both internally with key stakeholders and externally with the broader global development community, and include negative findings. This is because the goal is to share lessons to improve programming globally for greater development impact. Formal policies on sharing internally and externally are consistently followed and sharing is part of the program goal.			
Responsible	Guiding questions to the next level of the matrix 1. Is there an informal and/or formal policy on sharing data, results, and limitations of data?	Guiding questions to the next level of the matrix 2. Do you share any negative results internally with key stakeholders as well as externally with the global development community to share lessons learned and prevent others from making the same mistakes? Consider a Fail Festival. 3. Are informal and formal policies on sharing data and findings consistently followed?	There are no further questions.			
nsible – ect 3	MERL activities and processes do not check for unintended results or risks.	MERL activities and processes inconsistently check for unintended result or risks. D-MERL activities do not provoke inquiry when these results or risks are found.	MERL activities and processes consistently check for unintended results or risks. Actions to understand why these occurred as well as how to mitigate them are carried out.			
Responsible Aspect 3	Guiding questions to the next level of the matrix 1. Are you identifying potential negative outcomes or risks based on the program's theory of change and/or speaking to stakeholders to learn of the same?	Guiding questions to the next level of the matrix 2. Do you consistently collect data for unintended negative outcomes or risks from the program?	4. In the event that you found unintended negative outcomes or risks of the program, have you taken actions to mitigate these negative outcomes?			

Level 1: Unbalanced application of the principle		Level 2: Emergent application of the principle	Level 3: Balanced application of the principle				
Trustworthy: D-MERL is trustworthy when it is conducted according to standards of rigor appropriate to context, constraints, and/or intended use of the data. Trustworthy D-MERL is internally and externally valid, reliable, and objective, ensuring integrity of both the process and results.							
Trustworthy – Aspect 1	The program lacks a theory of change. A theory of change is the structure that provides outputs, outcomes, and goals; the theory of change also includes assumptions, risks, and linkages between different program components.	The program has a theory of change but it is not informed by a literature review and/or by key stakeholders. There is limited ownership of the theory of change by key stakeholders.	The program has a theory of change that is informed by a literature review and ratified/approved by key stakeholders. The theory of change forms the foundational document on which MERL activities and processes are then developed. Staff revisit the theory of change with key stakeholders during learning activities such as pause and reflect sessions and when decisions are being made on proposed adaptations to program design.				
	Guiding questions to the next level of the matrix 1. Have you developed a theory of change that accounts for leadership's program objectives and impact goals?	Guiding questions to the next level of the matrix 2. Where possible, is your program's theory of change informed by literature that suggests linkages between outcomes and impacts? 3. Have you reviewed the theory of change including assumptions and risks with your relevant key stakeholders?	Guiding questions to achieve this level of the matrix 4. Have you incorporated key stakeholders' feedback to improve the accuracy of the theory of change? 5. Do you use-the theory of change as the foundational document to develop your other MERL activities and processes? 6. Do you review and revise your theory change including assumptions and risks regularly based on MERL data and insights gathered?				
Trustworthy – Aspect 3 Trustworthy – Aspect 2	MERL indicators are not defined for each level of the theory of change. No thought is given to the indicators or their associated metadata until implementation of MERL activities and processes.	MERL indicators are attached to some or all of the levels of the theory of change. Some indicators do not measure the concepts they aim to capture. Not much thought is given to the metadata relevant to the indicator, until during implementation of MERL activities and processes.	MERL indicators are rigorous: they are well-defined for each level of the theory of change and measure the concepts they aim to capture. All metadata related to the indicator is included in the documentation during the planning stage and is consistently followed during implementation of MERL activities.				
	Guiding questions to the next level of the matrix 1. Did you identify key performance indicators (KPIs) for each level of your theory of change to establish how the program will measure each result identified?	Guiding questions to the next level of the matrix 2. Have you identified the metadata associated with your KPIs before MERL implementation? 3. Do your KPIs follow the SMART criteria (Specific, Measurable, Attainable, Realistic, Time-bound)?	Guiding questions to achieve this level of the matrix 4. Are your KPIs and metadata documented?				
	MERL activities and processes cannot provide data that meets the five dimensions of data quality: validity, reliability, integrity, precision, and timeliness. The data produced cannot be used to answer questions on program objectives and impact goals set by leadership. There is very limited documentation available to assess gaps in data quality of MERL activities.	MERL activities and processes can provide data that meets some of the five dimensions of data quality. These data can be used to answer certain questions on program objectives and impact goals set by leadership. There is some documentation of MERL activities and processes to understand gaps in data quality and what program objectives and impact goals cannot be measured.	MERL activities and processes can provide data that meets all five dimensions of data quality. These data can answer questions to all program objectives and impact goals set by leadership. MERL activities and processes are documented.				
	Guiding questions to the next level of the matrix 1. Are your MERL activities and processes responsive to some of the dimensions of data quality?	Guiding questions to the next level of the matrix 2. Are your MERL activities and processes providing you with data that meets all five dimensions of data quality?	Guiding questions to achieve this level of the matrix 3. Is your MERL analysis plan fully documented including the limitations of this plan?				

	L	evel 1: Unbalanced application of the principle	Level 2: Emergent application of the principle	Level 3: Balanced application of the principle	
	Tru	ustworthy: D-MERL is trustworthy when it is conducted according to standards of rigor appropriate to context, constraints, and/or intended use of the data. Trustworthy D-MERL is internally and externally valid, reliable, and objective, ensuring integrity of both the process and results.			
Trustworthy – Aspect 4	– Aspect 4	There is no formal data quality assessment system and no attention is given to it by leadership. There is no support given to MERL staff to implement data quality assessment procedures.	There is no formal data quality assessment system though some token attention is given to it by leadership. There is limited support given to MERL teams to implement data quality assessment procedures resulting in ad-hoc data quality assessment activities being conducted.	A formal data quality assessment system exists. MERL staff receive information, training, technology, and empowerment to implement the data quality assessment activities. Leadership pays attention to the results. Actions such as adaptations to the MERL design are taken to rectify all the issues identified through this assessment.	
	Irustworthy	Guiding questions to the next level of the matrix 1. Are you conducting any data quality assessments? Consider this checklist.	Guiding questions to the next level of the matrix 2. Have you developed a data quality assurance plan? Have you identified the standards of integrity and quality that you want to apply to your data? 3. Have your relevant staff been trained in the implementation of the data quality assessment procedures?	Guiding questions to achieve this level of the matrix 4. Have you identified activities to address any negative findings from the data quality assurance check? 5. Have you implemented activities to rectify any negative findings from the data quality assurance check been implemented?	

ACKNOWLEDGEMENTS

The authors wish to thank Innovations for Poverty Action and Search for Common Ground for their insights in early versions of the Maturity Matrix, as well as various colleagues at all of the consortium organizations who provided feedback on the matrix. The authors further extend their gratitude to USAID for their time and effort in providing feedback to draft versions of the document. The authors also wish to thank session attendees at the MERL Tech Washington September 2018 Conference and the ANDE Metrics Conference June 2018 for their feedback on ways to increase the usability of the matrix.









