



Guide to Measuring Implementation: The 3Us – Usability, Usefulness, Utilization

As the evaluation technical assistance partner for two portfolios of investments in Middle Years Math and Secondary Writing—funded by the Bill & Melinda Gates Foundation—Mathematica developed a set of tools to support grantees through a measurement and evaluation (M&E) process. The goal of these portfolios is to develop, refine, and scale evidence-based solutions (programs, products, or practices) that demonstrate success in improving educational outcomes for students who are Black, Latino, and/or experiencing poverty (the priority communities for the grants).

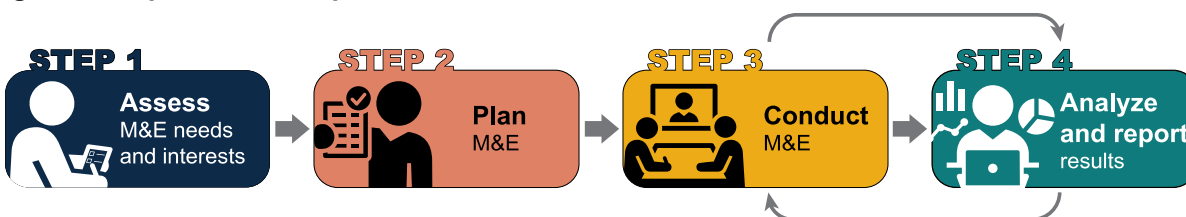
To expand the reach of this work and promote the adoption of evidence-based solutions more broadly, Mathematica has adapted the suite of M&E tools for a broader set of users—organizations implementing solutions, funders, research partners, and other stakeholders. The tools are designed to help users implement the M&E process in their own work. Organizations or individuals may choose to use one tool or the full set and can adapt the tools to their needs.

What is the M&E process?

The M&E process is designed to promote rapid innovation and scaling of promising solutions through generating timely and actionable evidence about what works for whom, and in what context. It uses an iterative approach to evidence building, in which the focus and design of the research is aligned to the solution's phase of development. Checkpoints are built in throughout to encourage users to pause and reflect on what has been learned so far and to refine the solution and the M&E plan as needed. More information on this iterative approach to evidence building and the phases of development is available [here](#).

In each evidence-building phase, the M&E process has four key steps (Figure 1). In Step 1, organizations articulate their M&E goals, interests, and needs. In Step 2, they develop a plan—including research questions—to guide the M&E work. In Step 3, they execute the M&E plan designed in Step 2, including collecting the data needed to address each research question. Finally, in Step 4, organizations analyze the data they collected in Step 3, determine next steps, and report the findings.

Figure 1. Steps in the M&E process



The M&E process centers on equity and community voice as a core principle. Across all steps, organizations are encouraged to meaningfully engage community stakeholders to ensure the solution builds on the strengths and assets and meets the needs of the community it is intended to serve.

Who should use this guide to measuring usability, usefulness, and utilization?

This guide is designed for organizations implementing solutions with support from a research partner. The research partner could be staff within the organization who have expertise in evaluation, or an external organization, technical assistance provider, or consultant.

What is this guide to measuring usability, usefulness, and utilization?

Usability, usefulness, and utilization are three key indicators of success in early implementation of solutions designed to improve outcomes.

- **Usability** refers to how easy or hard it is to implement and use the solution.
- **Usefulness** refers to users' perceptions of the solution's ability to meet their needs or provide advantages over alternative solutions.
- **Utilization** indicates the extent to which the solution is used, as measured by attendance rates, time spent using the solution, or other metrics.

Understanding how solutions perform in these areas is critical for implementing organizations and their funders, especially during development and early implementation. If solutions face challenges in these areas, organizations should adjust them to improve performance before attempting to assess the effectiveness of the solution, and before replicating and scaling the solution.

This document provides guidance on when and how to measure usability, usefulness, and utilization, and what implementing organizations can do with the results.

Acknowledgments

Mikia Manley, Virginia Knechtel, Kate Place, Andrew Gothro, and Ryan Ruggiero developed this guide. Gregory Chojnacki, Alex Resch, Megan Shoji, Mathematica technical assistance liaisons, and grantee partners provided feedback on the content. Sheena Flowers provided design and production support, and Jennifer Brown provided editorial support. This publication was prepared for the Bill & Melinda Gates Foundation. The findings and conclusions contained within are those of the authors and do not necessarily reflect positions or policies of the Bill & Melinda Gates Foundation.

Guide to Measuring Implementation: The 3Us – Usability, Usefulness, and Utilization

Overview

Usability, usefulness, and utilization are important indicators of successful program implementation. These indicators help organizations and funders understand whether their solutions meet the needs of users in the priority community and are used enough to be evaluated and potentially scaled. To understand how their solution performs in these areas, organizations or their research partners should measure these indicators at various points during solution development and implementation. Organizations should work with their research partner and get input from stakeholders in their priority community to determine which of these constructs are important to measure at each stage and how to measure them. The value of measuring each indicator will depend on the specific implementation context and may change across different stages of implementation.

When to measure usability, usefulness, and utilization

Table 1 provides an example of the different indicators that an organization might measure at different stages of a solution’s implementation.

Table 1. Indicators measured at different stages of implementation

Implementation stage	Usability	Usefulness	Utilization
Developing core components of solution	X	X	
Monitoring initial solution rollout	X		X
Examining association between solution use and improved outcomes			X
Scaling the solution	X	X	X

- Developing core components of solution.** In the initial design phase, an implementing organization would focus first on **usability and usefulness**. Investigating usability and usefulness would help the organization identify and improve core components of the solution and align them to the needs of the priority community. This would help an organization design a solution that stakeholders are motivated to use and can implement or navigate easily.
- Monitoring initial solution rollout.** When an organization begins to pilot its solution with the priority community, it might focus on measuring on **usability and utilization**. Measuring usability during implementation (in contrast to the controlled setting of the design phase) would help the organization determine whether it needs to make additional changes to the solution components or implementation approach. Measuring utilization would help the organization test whether the solution is being used as intended, a necessary precursor to changing outcomes.
- Examining association between solution use and improved outcomes.** If an organization is assessing whether the solution is associated with improved student outcomes, it might focus on **utilization**. This would enable the organization or research partner to analyze how the solution dosage is associated with the observed change in outcomes.

- **Scaling the solution.** If an organization is seeking to scale the solution, it might continue to focus on **utilization**. This would help confirm that users continue to use the solution as intended when it is made available more broadly. This is particularly relevant if organizations are scaling the solution to new contexts or populations. Organizations may also choose to conduct **usefulness** or **usability** testing with students and teachers who are new to using the solution to see whether new users have the same experiences with the solution as the users with whom it was initially tested.

How to measure usability, usefulness, and utilization

The specific methods used when measuring usability, usefulness, and utilization will vary according to the organization's context and circumstances and the implementation stage. Organizations or their research partners may use survey instruments; qualitative data from observations, interviews, or focus groups; attendance or usage records; or a combination of these. At the early stages of a solution's development, an organization may gather feedback from only a few users, while at later stages, an organization may wish to administer a survey or track attendance of everyone participating in the solution.

Organizations, research partners, or funders may choose to set targets before collecting data to set benchmarks for what they would consider success, given the context and stage of development. These targets could take several forms. **Outcome targets**, which assess the effect of the solution on outcomes of interest, might include expectations for the percentage of users who report finding the solution usable or useful, minimum average usability or usefulness ratings, enrollment or attendance rates, or average user time spent on solution components. Implementing organizations or their research partners might also articulate desired qualitative findings or specify in advance what actions they intend to take in response to the information collected. We refer to these as **process targets**. See the [Guide for Setting Targets](#) for additional information.

Definitions and examples

Table 2 provides more detailed descriptions of usability, usefulness, and utilization, along with examples of the types of measures organizations could use, the types of findings they might expect, and what types of targets they might set for each.

Table 2. Examples of usability, usefulness, and utilization in practice

Construct	Definition	Potential measures	Example finding	Target setting
Usability	<p>For technology solutions: How easy or hard it is to use the solution itself (such as whether an interface was easy to navigate) and how that might vary across different users (such as teachers and students).</p> <p>For both technology and non-technology solutions: How easy or hard it is to implement the solution (such as whether schools are able to incorporate tutoring sessions into the school day) and how that might vary across different users (such as teachers and students).</p>	<p>Observations: documenting how users explore, engage with, or implement the solution on their own (without guidance from observers)</p> <p>Error rate*: the proportion of incorrect interactions across all interactions. It is also important to document how users respond to errors.</p> <p>Time*: Amount of time it takes users to complete a discrete task</p> <p>Subjective feedback: surveys, interviews, or focus groups assessing the extent to which users report the solution is easy to use, how it could be easier to use, and how the solution does or does not fit with current practice. Subjective feedback can be influenced by beliefs, culture, and other context characteristics and should be paired with other measures of usability.</p>	<p>When navigating to the baseline assessment, users had an error rate of 30%. The organization will modify the assessment to address errors.</p> <p>It takes users 30 minutes, on average, to complete their intake assessment, more than the expected 15 minutes. The organization will provide clearer instructions and reduce the number of questions to decrease the time required for the assessment, with the ultimate goal of increasing assessment completion rates.</p> <p>37% of users report the solution is somewhat hard or hard to use, and a key driver is the navigation structure. The organization will incorporate feedback (which may be qualitative) into the solution design to make it easier to use.</p>	<p>It may be useful during development or initial implementation to set process targets focused on how the organization will identify and address usability issues. Organizations might set outcome targets during implementation or initial scaling to assess whether the solution’s usability meets desired levels.</p>

Guide to Measuring Implementation: The 3Us – Usability, Usefulness, Utilization

Construct	Definition	Potential measures	Example finding	Target setting
Usefulness	User perceptions of the solution’s ability to meet a participant’s need, provide an advantage over alternative solutions, or meet the objectives stated in the theory of change defined in the M&E plan .	Subjective feedback: surveys or qualitative user feedback sessions assessing the extent to which users report the solution is useful and why, and how it could be more useful.	70% of teachers report the solution meets their needs better than an alternative solution, falling below the organization’s 80% target. The organization will review current alternatives and modify the solution to add value compared to those alternatives.	It may be useful during development or initial implementation to set process targets describing how the organization will identify and act on opportunities to increase the usefulness of their solution. During early implementation or initial scaling, organizations might set outcome targets to assess whether the solution’s usefulness meets desired levels.
Utilization	The rate at which users take up the solution or the extent to which users use the solution.	<p>Take-up rate: percentage of individuals who use the solution after being offered access to the solution. When determining whether someone is “using” the solution, base the definition of usage on meeting a specific usage threshold.</p> <p>Usage: average time users spend on the platform or engaged with the solution (defined as minutes, hours, or logins) or attendance rates</p> <p>Completion rates: percentage of individuals who complete the full solution or specific tasks. In particular, organizations may want to look at the completion of tasks considered to be core components or look at how completion of certain tasks is correlated with overall utilization or solution completion.</p>	<p>55% of those offered the solution created accounts. The organization will reduce the registration fee to increase take-up.</p> <p>Students attend, on average, 80% of their tutoring sessions. The organization will adapt its scheduling approach to try to increase the attendance rate.</p> <p>60% of students complete the intake assessment, a critical component for determining which supports they receive in the program. The organization will send reminders to encourage students to complete the assessment.</p>	<p>During solution development, it may be useful to set process targets describing how the organization will identify and act on opportunities to improve solution utilization.</p> <p>During early implementation or initial scaling, organizations might set outcome targets to measure whether take-up or usage rates are high enough to improve student outcomes.</p>

* These measures may be most relevant for technology-based solutions.

Additional Measurement Guidance

1. Observations

Observations allow organizations to see how users explore, engage with, or implement the solution. They can be helpful for understanding where more guidance might be needed, or how the solution can be used in ways the solution designers did not anticipate. Organizations can document and summarize observations qualitatively. Early observation data can help define errors to be used in subsequent testing. When feasible, organizations should develop an observation protocol to help observers look for and document key information.

2. Errors

When measuring errors, organizations should determine a set of tasks and interactions that users must complete to fully engage with the solution. For example, users might need to register for an account, take an intake assessment, complete assignments, and request feedback in an online platform. To measure the usability of the solution, organizations may choose to observe users completing these tasks and record the ratio of correct to incorrect interactions. This is a useful approach if organizations are trying to determine how they can make their solution easier to navigate and is important for assessing the usability of core components. When recording errors, organizations should also record how users respond to errors. For example, are they able to troubleshoot on their own or do they get stuck? These findings will inform changes to the solution design or to user guidance.

3. Time

To measure the usability of the solution, organizations may choose to observe users completing these tasks and record how long it takes. This is a useful approach if organizations are trying to determine how they can make their solution easier to understand or navigate. If tasks are taking longer than expected, this might mean the directions are unclear or the task is too difficult. For example, if the organization expects the intake assessment to take 15 minutes on average, but it actually takes 30 minutes on average, the organization may need to shorten the assessment, modify the language in the questions, modify the instructions, or some combination of these approaches. Organizations might choose to seek qualitative feedback from users to understand which tasks felt like they took too long. This will provide insight into the value of reducing the time to complete a task.

References

Buzhardt, J., & Heitzman-Powell, L. (2005). Stop blaming the teachers: The role of usability testing in bridging the gap between educators and technology. *Electronic Journal for the Integration of Technology in Education*, 4(1), 13–29.

Nielsen, J. (2000). *Why you only need to test with 5 users*. Nielsen Norman Group.
<https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>