



HOW EVALUATION RUBRICS CAN HELP TRACE PROGRESS AND RESULTS IN COMPLEX PROGRAMS

Malene Soenderskov and Savi Mull

HOW EVALUATION RUBRICS CAN HELP TRACE PROGRESS AND RESULTS IN COMPLEX PROGRAMS

The complexity of international development work has increased over the past 25 years with a stronger effort to change the very systems that contribute to poverty and environmental damage. Yet our monitoring and evaluation (M&E) practices often fail to support these efforts. This article describes how evaluation rubrics can help planners and evaluators shift from rigid, cause-and-effect models to more dynamic, learning-oriented approaches that accommodate complexity, foster participation, and support adaptive decision-making in systems change interventions.

Malene Soenderskov and Savi Mull are measurement and evaluation experts that have experience in using rubrics with philanthropies, donors and partners.

1. From International Aid Projects to Systemic Change

Over the past 25 years, international development aid has shifted focus from isolated projects to addressing the structural, political, economic, and social dynamics that contribute to poverty, climate change, and environmental destruction. Thirty years ago, it was sufficient to provide direct humanitarian assistance in areas such as health, education, water, and sanitation. Today, country programs and framework agreements emphasize understanding and transforming the economic and political ‘ecosystems’ that perpetuate poverty. M&E

practices have not evolved accordingly, however.

2. The ‘Good Old Days’ of Project Support

In the 1990s and early 2000s, development projects were often bound by time and geography with measurable results in education, healthcare, water supply, and agriculture. Children were enrolled in school, vaccinations were administered, and agricultural output increased.

Projects followed a linear logic, assuming that successful experiences from one project could be transferred to the next, and that inputs would automatically lead to outputs and desired outcomes. This approach was neatly captured in a logframe, with goals set within three to five years.

Success was measured using quantitative indicators. These were easy to measure in terms of investments in projects, and it was relatively straightforward to document a project’s contribution to observed results—if it was not competing with other similar initiatives in the same location.

3. From Projects to Programs and Systemic Change – What’s Different?

It is much more challenging to measure change - let alone success in country programs and framework agreements that aim to drive fundamental, long-term change in political and economic ‘ecosystems’ that sustain poverty and marginalization.

¹ <https://realevaluation.com/what-are-rubrics/>

² <https://www.laudesfoundation.org/how-we-work/measuring-with-rubrics/>

HOW EVALUATION RUBRICS CAN HELP TRACE PROGRESS AND RESULTS IN COMPLEX PROGRAMS

Systemic change takes time, yet program cycles remain limited to 4-5 years. Measuring success through quantifiable indicators is difficult when today's programs focus on shifting power dynamics, changing narratives, and transforming relationships.

Systems involve multiple actors and factors that interact unpredictably, making it difficult to establish clear cause-and-effect relationships between inputs, outputs, and outcomes. Moreover, documenting and isolating the impact of a single program is nearly impossible in a context where multiple actors are driving change in different directions.

Nevertheless, many M&E systems still assume a project-based logic, where donors require results frameworks with clearly defined causal links between inputs, outputs, and outcomes. Programs are still designed with the expectation that we can track a positive change over time compared to doing nothing. However, M&E frameworks often fail to acknowledge that maintaining the status quo can be a success—and that preventing further deterioration of a system may be just as valuable as initiating change.

If our M&E does not account for these complexities, reporting risks becoming a mechanical description of short-term results, aligned with the logframe, rather than a meaningful reflection on whether

and how our work is contributing to broader systemic shifts, what adjustments may be needed when moving forward and which other approaches or partnerships may be needed for increased effectiveness and impact.

4, How Evaluation Rubrics can help overcome the challenges

'Systems' usually consist of power dynamics, relations between the system's actors, values, attitudes and perceptions that can be influenced by policies, competing discourses, changing funding streams or incentives that direct actors' behaviour within the system.

While changes and progress in these factors are hard to measure by a single, quantitative indicator, evaluation rubrics can help track progress and assess changes in systems, the 'value' of interventions over time and facilitate a discussion about how organisations may have contributed to these changes and 'what they can do next'.

¹ <https://realevaluation.com/what-are-rubrics/>

² <https://www.julianking.co.nz/vfi/subjective/>

HOW EVALUATION RUBRICS CAN HELP TRACE PROGRESS AND RESULTS IN COMPLEX PROGRAMS

What is an evaluation rubric?

An evaluation rubric is a structured framework used to assess and measure the status, based on predefined criteria and rating scales as in table one below. The evaluative rubrics framework for systems change was pioneered by Dr. Jane Davidson and Dr. Thomaz Chianca - and more recently discussed as a methodology that reduces subjectivity by Dr. Julian King. Rubrics provides a systematic and transparent way to evaluate a synthesis of qualitative and quantitative aspects of a project, program, or intervention.

Key Components of an Evaluation Rubric:

- 1. Criteria** – The specific and most valuable aspects being evaluated (e.g., shift in power dynamics, changes in narratives, as in table one below).
- 2. Status and levels of Progress** – A scale describing status or different levels of what good looks like (from unsupportive to supportive in table one).
- 3. Descriptors** – Clear explanations of what each status level looks like for each criterion.
- 4. Scoring Mechanism** – A way to collect data and assign values or ratings to progress levels. These can be both quantitative and qualitative or a synthesis of the two.

The following sections describe 3 ways evaluation rubrics can facilitate monitoring and learning in complex programming where a few, quantitative indicators are unable to grasp the width and depth of changes that might occur over the course of an intervention.

Criterion	Harmful	Emerging	Partly Conducive	Conducive	Thrivable
Shift in power dynamics	No evidence of shifting power or decision-making processes.	Decisions made by a small and inaccessible power elite. Most decisions made by a small group of decision-makers.	Some stakeholders report increased participation from time to time.	Increased representation and decision-making. Marginalized groups are usually engaged in decision-making, yet final decisions rest with a small group.	Power is fully decentralized, and decision-making is democratized, with marginalized communities leading governance structures and self-determined development processes.
Changes in narratives and discourse	Dominant narratives discredit and downplay the severity of the issues/problems addressed by our work.	Solutions-driven, evidence-based, and imaginative narratives have gained traction among the most progressive decision-makers and stakeholders.	Increasingly clear narratives are starting to drive a substantial number of decision-makers' sense of responsibility, motivation, and agency to deliver solutions.	Clear and compelling narratives effectively drive most key decision-makers' sense of responsibility, motivation, and agency to deliver – and hold each other accountable.	Narratives are not only solution-driven but also visionary and transformative, fostering a cultural shift where regenerative and just solutions become the default mindset in society.

HOW EVALUATION RUBRICS CAN HELP TRACE PROGRESS AND RESULTS IN COMPLEX PROGRAMS

4.1 Beyond quantitative indicators – assessing the uncountable

Most systemic changes – such as shifts in norms, attitudes, and policies – are difficult to measure with traditional M&E tools. Rubrics help evaluators and program managers assess such changes by breaking them into observable behaviours, conditions or performance standards, as in table one, providing a structured way to assess and measure progress based on predefined criteria and standards. These behaviours are also to show up as embedded changes and not one-time single efforts. For example, brands that sign up to pay premiums to organic farmers for the first year (project orientation) do so systematically as a sustained practice.

As such, rubrics can help provide an evaluative description of what ‘performance’ or ‘quality’ looks like, in areas that are hard to capture with a single indicator. Instead, rubrics allow evaluators to blend subjective and objective data to figure out, which performance standard best suits or describes the work. They can do so by combining narratives and qualitative assessments with quantitative data and structured scoring systems. This mixed-methods approach ensures that contextual insights complement numerical ratings and make the assessment more holistic and attuned to system-level changes.

4.2 Capturing Complexity and Systems Change

Systems change is rarely immediate or predictable. Yet with an evaluation rubric evaluators and program managers can track shifts over time, rather than just predefined outputs or outcomes – as illustrated in the textbox below. And because rubrics focus on qualitative judgment rather than fixed targets and indicators (as in a log frame), they can better capture emergent, unpredictable and nonlinear changes in complex systems.

Rubrics provide gradual, qualitative scales (e.g., from "no progress" to "transformational change"), that allows us to assess progress along a continuum rather than through rigid pass/fail judgments that are often associated with predefined targets and indicators. This allows for continuous learning and adaptation, as evaluators and program teams can use them in iterative cycles to adjust strategies, improving responsiveness to changing contexts and lessons learnt.

Evaluative rubrics have been the cornerstone of [Laudes Foundation’s measurement and learning approach](#) since 2021 and help assess contribution to systems change towards a green, fair and inclusive economy.

HOW EVALUATION RUBRICS CAN HELP TRACE PROGRESS AND RESULTS IN COMPLEX PROGRAMS

Using rubrics to determine status and decide on next steps

To determine progress of its child’s rights intervention in several countries, an international CSO developed and used tailored evaluation rubrics to assess status and performance. Program staff used secondary literature, semi-structured interviews with decision makers, focus groups with direct beneficiaries and available statistics to do so.

This process led to the conclusion that the intervention had contributed to move the situation from ‘harmful’ to ‘emerging’ on two criteria: Shifts in power dynamics and changes in narratives and discourse. Through the data collection and the barriers and successes that interviews with program stakeholders revealed, program staff further gained an understanding of ‘what it would take’ of the program to push the situation towards a more conducive situation.

Criterion	Harmful	Emerging	Partly Conducive	Conducive	Thrivable
Shift in power dynamics	Baseline →	Assessment Year 2	Milestone year 3-4		
	No evidence of shifting power or decision-making processes.	Decisions made by a small and inaccessible power elite. Most decisions made by a small group of decision-makers.	Some stakeholders report increased participation from time to time.	Increased representation and decision-making. Marginalized groups are usually engaged in decision-making, yet final decisions rest with a small group.	Power is fully decentralized, and decision-making is democratized, with marginalized communities leading governance structures and self-determined development processes.
Changes in narratives and discourse	Baseline →	Assessment Year 2	Milestone year 3-4		
	Dominant narratives discredit and downplay the severity of the issues/problems addressed by our work.	Solutions-driven, evidence-based, and imaginative narratives have gained traction among the most progressive decision-makers and stakeholders.	Increasingly clear narratives are starting to drive a substantial number of decision-makers' sense of responsibility, motivation, and agency to deliver solutions.	Clear and compelling narratives effectively drive most key decision-makers' sense of responsibility, motivation, and agency to deliver – and hold each other accountable.	Narratives are not only solution-driven but also visionary and transformative, fostering a cultural shift where regenerative and just solutions become the default mindset in society.

HOW EVALUATION RUBRICS CAN HELP TRACE PROGRESS AND RESULTS IN COMPLEX PROGRAMS

4.2 Enabling Participatory Evaluation and Multiple Perspectives

Last, but not least, evaluation rubrics allow us to incorporate context, multiple perspectives, and complexity into decision-making. If used in co-created processes, as in the case in [textbox 1](#), assessments of progress and performance becomes more inclusive, ensuring that the criteria for success reflect the perspectives of communities, program implementers and other stakeholders, rather than external donor-imposed metrics. As such, evaluation rubrics can facilitate reflection and sensemaking, allow for deeper discussions about what is working, what is not, and why, fostering a culture of learning over compliance.

Want to know more?

Drop us an email on mail@strategyhouse.dk

Malene Soenderskov and Savi Mull are measurement and evaluation experts that have experience in using rubrics with philanthropies, donors and partners.

www.strategyhouse.dk

Who can use rubrics? Who are they for?

Evaluative rubrics can be used by organisations for their strategy, for programmes, and projects. They define what good looks like because they use a rating scale and descriptors for each scale.

As Julian King points out, rubrics can be structured in 'unlimited ways' - including holistic, generic, and analytic approaches - and these various approaches can be implemented based on the specific unit or universe of measurement required. Rubrics are ideally penned before change starts to take place. Evaluators, too, can recommend the use of rubrics for complex programmes that focus on non-linear, emergent changes within differing, operating contexts.

Organisations/programs can use rubrics:

- At start of strategy development, strategic review processes
- At defining programme strategy development or a review process
- During evaluations

Evaluators can use rubrics:

- In proposals if programmes are complex and focused on systems change
- During baseline, midline and endline processes
- For Developmental evaluations
- To support a multi-method evaluative approach