



# The ELMA Philanthropies

## Evidence & Learning Principles and Practices

June Lee  
Director, Evidence & Learning

---

### Purpose and Audience

The Evidence and Learning (E&L) function at ELMA is guided by a set of principles that aim to contribute to effective, impactful philanthropic investments for The ELMA Group of Foundations. These principles have been informed by multiple sources: ELMA's overarching approach to philanthropy, formal and informal input from staff (program and non-program) and board members,<sup>1</sup> and an extensive review of other foundations' M&E thinking and practice.

This document is intended for Philanthropies staff and our evaluation partners. It codifies a set of shared principles and guidance for evidence and learning, with the ultimate goal of yielding objective data that assess the social returns on our investment, and can contribute to evidence bases that may inform broader policy, practice, and funding. It does not explore all the technical and non-technical considerations that should go into an evaluation.

### I. Evidence can take many forms

The purpose of evidence is to assess and understand the progress or outcomes of ELMA's investments. It can include data from monitoring, from various types of evaluations, investments' measures of success, and grants management. All are used to generate evidence—of varying depth—that grant investments are on the path to success. Typically, evaluation and monitoring are complementary (see Table 1). Monitoring is the ongoing tracking of programs to ensure that they stay on course, and usually refers to routine activities conducted by implementers/grantees. ELMA staff also actively monitor grant investments—through our interim reporting processes, formal and informal check-ins with grantees, site visits, etc.—to assess if they are on or off track in reaching their goals, and to provide additional support as necessary.

Evaluations can also take many forms; they are distinct from monitoring in that they aim to provide an objective, systematic, and independent appraisal of the program (e.g., through a deeper dive into elements of a program model, assessing theories of change). A range of methods can be used in evaluations, including experimental, quasi-experimental, process, and qualitative methodologies. They can focus on processes or outcomes, and answer broader questions about lessons learned. Evaluations are usually (but not always) conducted by an independent third party and can be commissioned by ELMA, a co-funder, or the grantee. ELMA M&E staff oversees ELMA-commissioned evaluations and in the case of co-funder or grantee-commissioned evaluations, provide input to shape them. Data from both monitoring and evaluation are evidence and can be used to validate an investment's measures of success.

The intended user of evidence can be the implementer/grantee; ELMA; governments to make, fund, or fundraise for a policy change; other funders when making funding decisions related to a specific organization or intervention; other implementers who want to test or adapt an intervention; or the wider world (e.g., academic research audiences). The type of evidence and approach can depend heavily on its primary user.

---

<sup>1</sup> This included a survey of board, program, and non-program staff involved in grant recommendations was fielded in March 2018 to understand staff and board perspectives on the priorities and gaps in ELMA's current M&E practice.

Table I. Monitoring vs. evaluation

	Monitoring	Evaluation
<b>Purpose</b>	<ul style="list-style-type: none"> <li>• Focus on implementation, comparing what is being implemented vs. what was planned.</li> <li>• Keep project on track; identify areas of underachievement; make course corrections</li> <li>• Provides lessons learned and offers recommendations</li> </ul>	<ul style="list-style-type: none"> <li>• Can focus on process and/or outcomes</li> <li>• Validates what was achieved</li> <li>• Examines how and why outcomes were (or were not) achieved</li> <li>• Examines issues such as best practices, adaptation of models, scale, sustainability</li> <li>• Provides lessons learned and offers recommendations</li> </ul>
<b>Timing</b>	Routine, regular, throughout the project	<ul style="list-style-type: none"> <li>• At significant points during or after the project</li> <li>• Selective and episodic</li> </ul>
<b>Scope</b>	<ul style="list-style-type: none"> <li>• Activities/inputs</li> <li>• Outputs</li> <li>• Outcomes</li> </ul>	<ul style="list-style-type: none"> <li>• Process, intermediate or final outcomes</li> <li>• Test underlying theory of change (or parts thereof) or dive deeper in elements of a program’s model</li> <li>• May try to understand causality</li> </ul>
<b>Sources/ Participants</b>	Internal staff or implementers	Internal and external: Evaluators, implementers, funders, stakeholders, partners
<b>Users</b>	Internal staff or implementers Funder	Internal and external: Implementers, funders, stakeholders, partners
<b>Examples</b>	Routine monitoring data on reach, training, convenings, etc.; administrative data	<ul style="list-style-type: none"> <li>• Needs assessment</li> <li>• Process/implementation evaluations</li> <li>• Outcome evaluations</li> </ul>

Table I describes the attributes typical of monitoring and evaluation, with some characteristics that are common to both (e.g., monitoring and process evaluations can yield similar information).

## 2. Evidence should be forward-looking and prioritize decision making and learning

ELMA engages in evaluation activities to assess the effectiveness of program strategies and specific investments, inform internal decisions, and inform and ideally influence decisions by grantees, government, co-funders, and other partners. Decision making and learning<sup>2</sup>—rather than accountability or policing—is the core focus of E&L and evidence should be linked to decision points so that they are timely and useful. A common pitfall in evaluations is that findings are designed and delivered too late to be useful. For example, an impact evaluation of a two-year program may be completed at the end of the third year, and ELMA has had to make investment decisions in absence of the evaluation findings. Evaluations are of limited utility at that point, and intentions to “keep the results in mind for future investments” are often unfulfilled. In some cases, program outcomes are not fully realized until some time after the end of the grant, and outcomes data are not yet available for decision-making, in which case information from other ongoing measurement efforts will help to validate that the program is on the path to achieving impact and provide confidence in investment decisions.

<sup>2</sup> “Learning” in this case prioritizes learning on the part of ELMA, implementing partners, and/or policymakers, rather than academic learning.

Evidence also cannot inform decision making without a clear plan for its use. Therefore, evidence planning must happen early in the grant development and strategy-setting process, and planned in tandem with program design.

Once it is determined that a grant or cluster can benefit from external evaluation, the plans must be designed with a clear understanding of:

1. what decisions need to be made
2. when they need to be made
3. what specific evidence is needed to make the decisions
4. what ELMA would do differently if we had that evidence
5. who (internal and external) will use that evidence and how best to communicate it

If a program aims to influence policy or practice, additional considerations include:

6. who the audience is
7. how the evidence fits within a broader influence or advocacy strategy

The evidence planning tool in Appendix A is intended to facilitate these discussions.

It is important to note that evidence is one piece—but not the only piece—of information used in ELMA's decision-making.

### 3. We prioritize evidence for grant investments that are large and/or otherwise strategically important and where evidence has the potential to influence policy, funding, and practice

Formally evaluating every program in which ELMA invests is not a feasible use of time or resources. We keep several criteria in mind when thinking about whether a grant or cluster can benefit from an evaluation, evidence-generation, or learning support, what kind, and how. These considerations fall along the dimensions of importance and influence (see Fig. 1).

**Importance** indicates the strategic importance of the grant to ELMA, including:

1. expected outcome/impact
2. size of grant investment<sup>3</sup>
3. ELMA's long-term commitment to the strategy
4. a program's potential scalability
5. a program's potential to improve policy or practice

**Influence** denotes the potential for evidence to have influence (especially externally), such as:

1. potential learning (e.g., new and/or untested approaches; plausibility of grantee's theory of change; contribution to the evidence base)
2. urgency for course corrections or future funding decisions
3. potential leveraging (e.g., providing a basis for securing additional funding)
4. alignment with grantee and stakeholder/government concerns
5. a program's potential to improve policy or practice

---

<sup>3</sup> Not every strategically important grant is large, but large grants are by nature strategically important because of the substantial investment that ELMA has made.

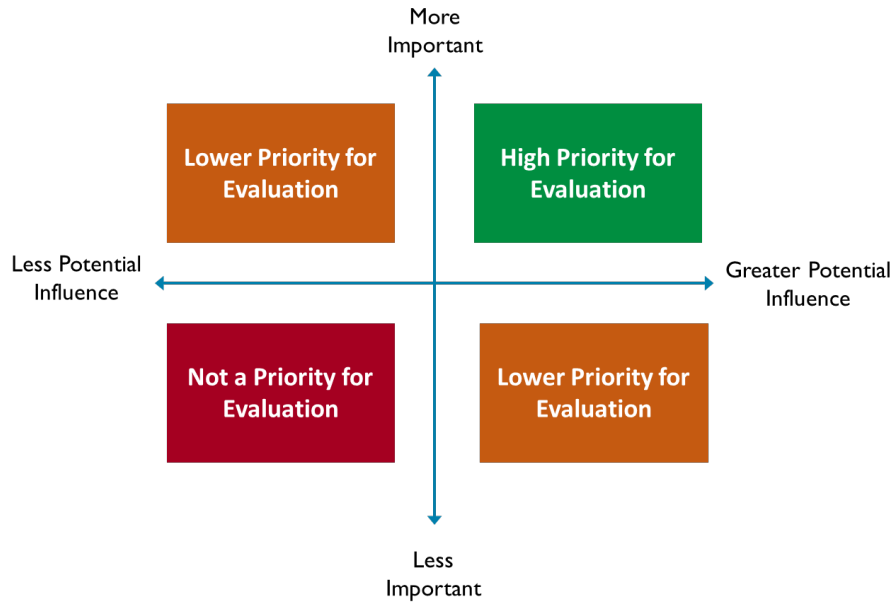


Figure 1. *Weighing importance and influence to prioritize evidence*

Not all criteria receive equal weight, and different criteria may be used to deliberate what to evaluate versus how intensely to evaluate. The more a grant or cluster meets the criteria above, the more important it is to engage in serious and thoughtful planning for evaluation. Appendix B contains a checklist of considerations to help prioritize evaluations.

#### 4. Plans for evidence-gathering should be fit for purpose and right-sized

##### **Fit for purpose: The purpose of the evidence determines its design.**

Evaluations should be fit for purpose. That is, the purpose of the evaluation should determine its design. (see Table 2). Different kinds of evaluations can help to distinguish between failures of theory (i.e., program strategies do not lead to the intended outcome because the Theory of Change is flawed) and failures of implementation (i.e., the theory may be correct, but resources and capacities were insufficient to implement it and produce results), barring unexpected external shocks/factors. Understanding when each type of evaluation is useful and what kind of information is needed will help to ensure evaluations serve their purpose and are designed accordingly. The evaluations described below can be used to inform a program before or during implementation, or assess processes and outcomes after the fact. The rigor and resources committed should be proportionate to a grant’s strategic importance and size.

Table 2. *Examples of different types of evidence*

Type of Evidence	Purpose	When?
<b>Needs assessment or landscape analysis</b>	Determines the need for the intervention among a specific population, and how/if the intervention can meet the need; assesses where and how a program should work	Developing new program or modifying a program for a new context/population

Type of Evidence	Purpose	When?
<b>Process/Implementation Evaluation</b>	Examines whether a program is implemented as designed among the target population.	The onset of a program and on an ongoing basis
	Identifies barriers and enabling factors to implementation. Enables early course correction	
	Often complemented by monitoring data	
	Look more deeply at issues that arise during the course of implementation	During or after program implementation
	Lessons from implementation process that guide continued movement to scale and/or inform other efforts	After implementation/at the end of the program
<b>Outcome evaluation</b>	Examines prospects for achieving outcomes, leading indicators, and changed assumptions	Early or during a program
	Measures medium- and long-term outcomes from a program to understand whether a program achieved its objectives, for whom, and how. Outcomes can happen long after the end of a grant.	During, at the end of, or after the end of a program
	Assess the strength of a causal link	
	Can include cost-effectiveness analyses	
<b>Sustainment assessment</b>	Understand how impact is sustained, and (where relevant) contribution to systemic change	After the end of a program

In addition to thinking about what type of evidence is necessary, we also consider the *level* of evidence needed. The standard of evidence should be proportional to the nature of the investment and the decisions or actions to be made. For example, a program that we hypothesize can or should be scaled widely or that has implications for practice and policy prescriptions likely requires an established, rigorous evidence base of its effectiveness and scalability. In fact, the evidence generated can contribute to more funding and effective scaling. Evidence for programs in a phase of testing and adaptation, on the other hand, may be oriented toward learning and establishing sound measurement practices rather than on rigorous proof of impact. Such decisions have to balance factors such as technical considerations, cost, time, operations, political considerations, etc. A consultative process that enables rich discussions that weigh different dimensions of rigor and practicality will help to achieve the optimal design. As mentioned earlier, the type of evidence and approach can also depend heavily on its primary user(s).

Different quantitative and qualitative evaluation methods are complements to improving understanding, and we believe in their judicious use. Some evaluation designs (such as randomized controlled trials) can be complex and costly and have many design requirements. When it is impossible to meet these requirements or where key aspects of evaluation design have to be compromised, resources can be put to better use by bolstering the quality of monitoring data.<sup>4</sup>

---

<sup>4</sup> Gugerty, M. K., & Karlan, D. (2018). *The Goldilocks challenge: Right-fit evidence for the social sector*. New York: Oxford University Press.

## 5. Evidence generation must be collaborative and consultative

For evidence to be useful, it must be designed in consultation with their primary users, both internal and external, to ensure that user perspectives and interpretations are heard, considered, and incorporated appropriately. Learning questions should also reflect information that decision makers need. Internally, ELMA M&E staff is committed to engaging with program teams throughout the process, including conceptualizing the learning/evidence questions, developing a plan, providing or reviewing suggestions for consultants, etc. Staff are actively involved in reviewing progress (including interim findings), and sharing and discussing results. It is important to acknowledge that such a participatory approach—important and often desired—requires considerable time, attention, and collaboration.

Externally, ELMA often engages with a broad range of stakeholders (e.g., government, co-funders, grantees, and stakeholders in the field at large), and involving appropriate key stakeholders in evidence generation is an important precursor to getting strong buy-in and incentivizing the uptake of learning. Stakeholder engagement, however, can be a double-edged sword, with the need to prioritize ELMA's goals in balance with different stakeholder interests, desires, and standards of evidence.

From a grantee's perspective, funder-commissioned external evaluations can be particularly fraught, especially given the inherent power imbalance in the funder-grantee relationship. Grantees can feel that evaluations have high stakes, and that future funding decision rests on an evaluation over which they have very little say. It is important that evaluation is understood as an organized decision-making and learning process that involves grantees and their knowledge and perspectives, rather than an exercise that is “done to” them. Clear communication is essential, including being clear about the decisions that the evaluation is intended to inform so that there is no ambiguity about an evaluation's purpose and that it does not cause undue anxiety or create false expectations. In ELMA's *Our Approach to Philanthropy* document, we describe the importance of investing in organizations that are evidence- and data-driven. Mutual buy-in to M&E is important, and very strong resistance to evaluation from the grantee may undermine the evaluation and the investment. It is also important to note that participating in an evaluation can require significant time commitment from grantees, and it is often necessary to resource them to do so.

## 6. We are committed to using and sharing evidence

### Creating time and space for internal learning

Amid full workloads and busy schedules, it is challenging to intentionally make the time and space for learning. Opportunities for such reflection—both about the program and the evidence itself—include brown bags, learning sessions, and evaluation presentations, briefs, or memos. We recommend three questions to elicit reflection: What? So what? Now what?<sup>5</sup> These questions are intended to strengthen the link between evidence and action:

**What?** entails an understanding of what happened on a grant or cluster of grants, gleaned through the evidence, the investment's Measures of Success, and program teams' expertise on the grant.

**So what?** examines why what happened was important, and what lessons—both specific to the intervention and/or context, as well as generalizable for the field at large—can be learned.

**Now what?** looks at future actions, such as what ELMA or grantees will do differently in the future, whether and how grantmaking /program strategies should change, and what ELMA will do with the evidence to influence action.

### Tailored communication is a critical to ensure external uptake of results

In addition to informing internal decision-making, evidence often has the goal of having external influence (e.g., crowding in funding, affecting policy or practice). Communicating and disseminating findings strategically are critical

---

<sup>5</sup> Human Systems Dynamics Institute. (n.d.). Adaptive Action. <https://www.hsdsinstitute.org/resources/adaptive-action.html>  
“What? So what? Now what?” broadly correspond to findings (what the evaluation found), conclusions (interpretations of the findings), and recommendations (what actions should be taken as a result of the conclusions), respectively.

to have the desired influence. Knowledge products that emerge from cannot be “one report fits all” and must be tailored to the interests and needs of varied audiences, with relevant products (e.g., summaries) made available and accessible for different consumers of the information. In addition, facilitated discussions of the results must take place in a timely manner so that they can be interpreted, understood, and acted on in the context of each stakeholders’ needs. As such, evidence generation should be accompanied by a communication plan (see Appendix C) that is developed in partnership with program teams and others.

Communication plans should consider whether it would be appropriate to share findings with the following partners:

**Grantees.** Evidence has the most immediate implications for grantees’ programs and are not only shared with, but co-interpreted, with grantees.

**Other partners.** Government partners, funding partners, communities, and/or participants, should receive a relevant summary (and if appropriate, the full report), participate in learning discussions, and be invited to provide input on avenues of communication.

**Program participants.** Program participants are often respondents in evaluations but hear nothing back after the they are completed. It may not always be possible or feasible to share the results with all participants, but we should consider the best way to share or give back so that they do not experience evaluations as a one-way extraction of information.

**Policymakers.** Policymakers are a unique audience and must consider a wide range of interests and issues when adopting or changing policy.<sup>6</sup> They use a broad range of evidence and information from a variety of sources, most of which may not be rigorous scientific studies. When ELMA reviews investments that have the ultimate goal of informing policy or practice by government, ELMA and grantees should have a plan for *using* evidence to advocate for and support government champions on specific policies or practices. While ELMA has built strong relationships with government and policymakers in many countries, using evidence to inform policy is a complex process,<sup>7,8</sup> and no single evaluation will shift policy on its own. It is important to understand what data are especially important for policymakers at which juncture (e.g., cost effectiveness, how to scale a program, acceptability of a program among constituents) and (where possible) include government on discussions around generating evidence, while acknowledging that evidence is a necessary but not sufficient condition for policy change.

**The public.** ELMA can use multiple platforms to reach broad audiences and the public with our learnings. Our website, conferences, publications, and blog posts are all part of ELMA’s communication strategy about evidence.

## Conclusion

Ultimately, ELMA’s Evidence and Learning function serves the goals of generating knowledge, making knowledge usable, and enabling a response to knowledge. These principles provide guidance for how ELMA aspires to approach its E&L practice, while recognizing that learning is in fact a messy and nonlinear process, and that ELMA itself will learn from it and continue to refine its approach.

---

<sup>6</sup> See, for example, Shaxson, L. (2014). Investing in evidence: Lessons from the UK Department for Environment, Food and Rural Affairs. [http://www.ksi-indonesia.org/files/1421384737\\$1\\$QBTM0U\\$.pdf](http://www.ksi-indonesia.org/files/1421384737$1$QBTM0U$.pdf)

<sup>7</sup> Dhaliwal, I., & Tulloch, C. (2012). From research to policy: Using evidence from impact evaluations to inform development policy. *Journal of Development Effectiveness*, 4:4, 515-536. DOI: [10.1080/19439342.2012.716857](https://doi.org/10.1080/19439342.2012.716857)

<sup>8</sup> Jones, H., Jones, N., Shaxson, L., & Walker, D. (2013). Knowledge, policy and power in international development: A practical framework for improving policy. <https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/8201.pdf>

## Appendix A: Evidence Planning Tool

The Evidence Planning Tool is a basis for conversations between program and E&L staff on evaluations. It should be used when evidence needs are first contemplated and is a reference for critical questions to answer. This will help to make sure that evaluations fit our E&L principles, and are sound and well utilized.

FOR EVIDENCE PLANNING:	
What decisions need to be made?	
When do they need to be made?	
What specific evidence/information is needed to make the decisions?	
What would ELMA do differently if we had the evidence?	
Who needs the evidence? (Internal)	
Who else needs the evidence? (External)	
What/Who are we trying to influence?	
Who will use the results for influence and learning?	
Who will make sure the learnings are acted on?	



## Appendix B: Should We Think About Rigorous Evaluation? Checklist

How do we know when we should consider a more rigorous evaluation? This checklist can help as a tool for decision making. The more check marks an investment has, the more likely E&L and program staff should consider a rigorous external evaluation.

### **Importance** (the strategic importance of the grant to ELMA)

- the expected outcome/impact is high
- large grant investment
- ELMA's has a long-term commitment to the strategy
- program's has high potential for scale
- program's has high potential to improve policy or practice

### **Influence** (the potential for evaluation findings to have influence, especially externally)

- potential learning from the evaluation is high (e.g., new and/or untested approaches; plausibility of grantee's theory of change; contribution to the evidence base)
- there is urgency for course corrections or future funding decisions
- potential for leveraging evaluation findings is high (e.g., providing a basis for securing additional funding)
- grantee and stakeholder/government concerns are well aligned
- program's has high potential to improve policy or practice

### Appendix C: Evidence Communication Planning Tool

The Evidence Communication Planning tool is a reference to help ELMA M&E and communications staff and evaluators to plan for and resource communications around evaluations.

<b>Knowledge Product</b> [Report, summary, presentation, meeting/workshop, etc.]	<b>Audience</b>	<b>Target Date</b>	<b>Lead Contributors</b>	<b>Dissemination Partners or Outlets</b>	<b>Resources Needed</b>