USAID
FROM THE AMERICAN PEOPLE

VERSION 2.0 | DECEMBER 2013

Discussion Note

Complexity-Aware Monitoring

Monitoring & Evaluation Series

USAID typically uses a specific approach to monitoring, referred to as performance monitoring. Performance monitoring uses *indicators* designed to measure *results* that contribute to broader country strategy results frameworks or project LogFrames. Annual (or semi-annual) review of country strategy (CDCS) performance data is intended to inform high-level decision making. At the project and activity level, monitoring is intended to inform implementation. Performance monitoring practice involves collecting *baseline data*, setting *targets*, and comparing actual figures to targets. (For more information on USAID's approach to performance monitoring, please see the Automated Directives System section 203.3.2).

Outside the Agency, the term "monitoring" may be used to describe a much broader array of practices with roots in diverse theoretical perspectives. For example, monitoring, for other organizations, does not necessarily involve results, indicators, baselines or targets.

This discussion note outlines general principles and promising approaches for monitoring complex aspects of USAID development assistance. Complexity-aware monitoring is distinct from performance monitoring as practiced in USAID and is intended to complement performance monitoring when used for complex aspects of projects and strategies. Complexity-aware monitoring may be considered "normal" monitoring by some working in other organizations or contexts. Nevertheless, consideration of these principles and approaches may strengthen practice.

This Discussion Note is designed to prompt inquiry and experimentation within USAID with new approaches and methods for monitoring complex aspects of development assistance. Rather than prescribe a single method or approach, this note highlights principles and methods used by development practitioners outside of USAID. Developed in consultation with outside experts in the principles and methods described and with USAID staffers who are already experimenting with new M&E methods, it is a starting point for USAID staff wishing to experiment with methods that suit some aspects of their portfolios better than performance monitoring as described in ADS 203. After a period of experimenting with and learning how these methods and approaches work in USAID's programming context, this Discussion Note may be turned into a How-To or Technical Note that will broaden USAID's M&E toolkit.

This publication was produced for review by the United States Agency for International Development. It was prepared by Heather Britt for DevTech Systems, Inc., under Contract No. AID-OAA-M-11-00026. The views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

WHEN TO USE COMPLEXITY-AWARE MONITORING

Complexity-aware monitoring is appropriate for aspects of strategies or projects where cause and effect relationships are poorly understood, thereby making it difficult to identify solutions and draft detailed implementation plans in advance. [1] Expected results may also require refinement and revision as strategies and projects unfold. Projects (or parts of projects) that rely heavily on adaptive management[2] to steer effectively in dynamic contexts, and projects that seek to influence social change or innovate to discover solutions are likely candidates for complexity-aware monitoring. On the other hand, projects that deliver services, or roll out, replicate, or scale up tried and true programming strategies are not generally a good match for these monitoring approaches.

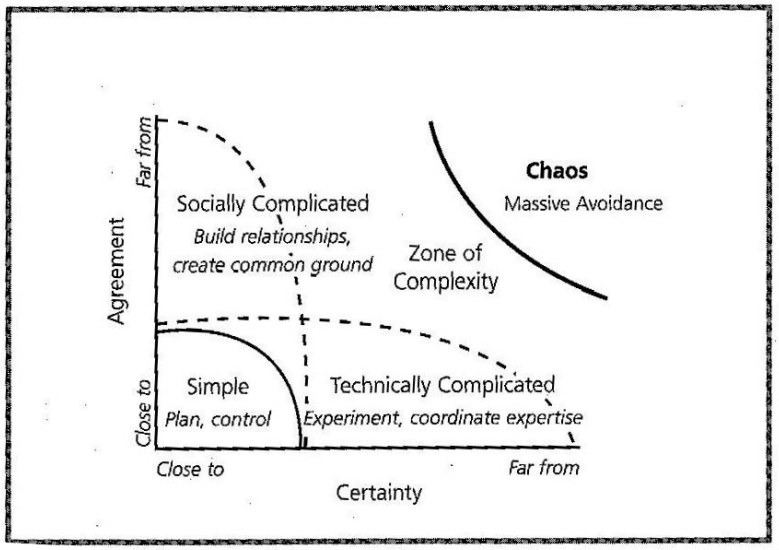

FIGURE 1: AGREEMENT AND CERTAINTY MATRIX
Source: Patton, M.Q. (2011). Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use. New York: The Guildford Press, p. 94.

Two questions can help identify complex aspects of a project or strategy:

- What is the degree of certainty about how to solve the problem?
- What is the degree of agreement among stakeholders about how to solve the problem?

According to the Agreement and Certainty Matrix[3] depicted in Figure 1, complex aspects of situations are distinguished from simple and complicated by both low certainty and low agreement. In situations of low certainty, even the experts are uncertain about the best way to achieve results. In low agreement, key stakeholders disagree about which results are desirable.

Those in the Agency working with the Cynefin framework,[4] depicted in Figure 2 (next page), recognize complexity when diverse elements interact with each other in unanticipated ways to create a new reality. Complex aspects of a situation cannot be known or predicted ahead of time; cause-effect relationships emerge

[1] In 2011, USAID revised its definition of project from a single intervention or implementing mechanism such as a contract or grant to "a set of planned and then executed interventions identified through a design process, which are together intended to achieve a defined development result, generally by solving an associated problem or challenge." (ADS 201.3.7) In the new USAID lexicon, projects are comprised of activities, "any mechanism or other interventions using program or operating expense funds below the project." Each activity is generally carried out by a single implementing partner who receives funds by an instrument or mechanism such as a contract, grant, or cooperative agreement. The term "activity" refers to what was formally considered a project, but it also may be used to refer to initiatives, such as strategic communications, carried out directly by the Agency to promote a project purpose. In this Note, the terms "programming" or "intervention" are used to refer to purposeful efforts to influence development objectives, regardless of whether they involve disbursement of funds.

[2] Adaptive management refers to a flexible management approach that enables projects to catalyze and respond to contextual change within standard implementing mechanisms.

[3] An early discussion of this matrix can be found in Zimmerman, B., Lindberg, C., & Plsek, P. (1998). Edgeware: Lessons from complexity science for health care leaders. Dallas, TX: VHA Inc. (page 141).

[4] Kurtz, C. F., & Snowden, D. J. (2003). The new dynamics of strategy: Sense-making in a complex and complicated world. *IBM Systems Journal*, 42(3), 462-483.

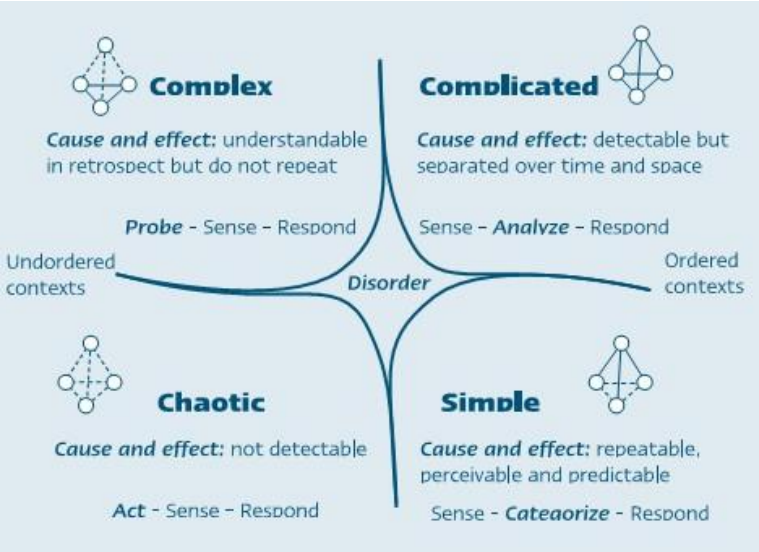

FIGURE 2: CYNEFIN FRAMEWORK
Source: Wageningen UR, 2012.

only retrospectively. Complexity responds well to probe-sense-respond management approaches, in which development practitioners experiment, gather information, and then act accordingly.

The Cynefin framework contrasts complex dynamics with simple, complicated and chaotic. In simple aspects, causal dynamics are well known. The right answer is common knowledge. Best practices have been identified. In complicated aspects of a situation, cause-effect relationships are knowable with additional expertise or time and energy to understand and measure. Experts would be expected to possess the relevant knowledge, and to be able to identify effective good practices. Alternatively, piloting, testing or experimenting would serve to unveil the cause-effect relationships. Chaotic aspects of a situation are ones in which there is so much turbulence that causal relations are not perceivable even after effects emerge and there is no time to investigate or figure out right answers. Those dealing with chaotic aspects must act quickly and decisively to reduce the turbulence.

Projects and their environments can have some simple aspects, some complicated aspects and some complex aspects; it is more useful to identify these individually than to attempt to classify a whole situation as either one or the other[5]. In general, social change and development contexts and programming tend to contain a mix of complicated and complex aspects. Thus, complexity is neither a special circumstance (such as conflict or transition), nor a blanket descriptor. Furthermore, judgments about complexity are ones about relative disagreement and uncertainty. Complicated aspects of a system may evolve to demonstrate complex dynamics or complex aspects may become complicated.[6] Also, aspects of complexity worth monitoring may cross boundaries of projects, activities, and contracting mechanisms.

[5] Rogers, P. (2011). Implications of complicated and complex characteristics for key tasks in evaluation. In K. Forss, M. Marra, & R. Schwartz (Eds.), *Evaluating the Complex: Attribution, Contribution, and Beyond* (p. 39). New Brunswick, New Jersey: Transaction Publishers.
6 Kurtz, C. F., & Snowden, D. J. (2003). The new dynamics of strategy: Sense-making in a complex and complicated world. *IBM Systems Journal*, 42(3), 462-483.

Let's take the example of a vaccination project (Table 1). Aspects of the project demonstrate simple, complicated, and complex dynamics. Vaccines work by introducing a modified version of a disease agent into the body, thereby stimulating the body's immune system to build up defenses against the pathogen. If a vaccinated individual encounters the disease agent again, the immune system will be able to ward off the disease. The cause and effect relationships are known and the effectiveness of vaccines is very high. Preventing common childhood diseases is generally agreed to be a worthwhile goal. Therefore, both certainty and agreement about how vaccinations work and the desirability of preventing childhood disease of the project is high. We can call this aspect of the project **simple**. Simple does not mean that a successful vaccination project will be easy.

TABLE 1: CERTAINTY & AGREEMENT IN SIMPLE, COMPLICATED & COMPLEX ASPECTS OF PROJECTS

Simple	High certainty, high agreement	Stimulating immunity through vaccines
Complicated (technically)	Low certainty, high agreement	Delivering vaccines to remote populations through a weak health system; population values vaccines
Complicated (socially)	High certainty, low agreement	Delivering vaccines to resistant communities through a well-resourced health system
Complex	Low certainty, low agreement	Delivering vaccines to resistant communities through a weak health system

Other aspects of the project, such as logistics, may not be simple. Implementing a vaccination project can be challenging in a poor country without a strong health system, even if the population is generally supportive. Many vaccines require cold storage up until they are administered. Rural populations are difficult to access without an adequate transportation network. When facing logistical challenges but working in a population that shares common values of protecting the health of children through vaccination, delivery of the project can be said to be **technically complicated**.

In some situations, the value of the vaccination effort itself may be in question. Recently, the US has witnessed the rise of an anti-vaccination movement that claims a link between vaccinations and autism. Other studies have linked low rates of vaccination to lack of trust in medical workers among certain populations in the US. Both of these populations exhibit low agreement with the project's theory of change and underscore the need for cultural sensitivity in vaccination campaigns. The US health system makes vaccinations accessible in most parts of the country, but low agreement about the value of vaccinations makes reaching specific populations **socially complicated**.

Different interpretations of the value of vaccinations can be found in countries as diverse as Chad and the US, Australia and Pakistan. When working in a situation in which logistical challenges lower the certainty and cultural issues lower the agreement, delivery of the project is **complex**.

Performance monitoring as currently practiced within the Agency relies on predictive practices built on known or hypothesized cause and effect relationships. For this reason, they are better suited for the simple and complicated aspects of a strategy or project. Complexity-aware methods, which can monitor dynamic and emergent aspects of projects and strategies, can thus complement and enrich performance monitoring.

PRINCIPLES OF COMPLEXITY-AWARE MONITORING

The application of the three key principles listed below can be invaluable for monitoring the emergent and dynamic aspects of strategies and projects. When you apply these principles to your own situation, you may discover new monitoring solutions.

1. Synchronize monitoring with the pace of change
2. Attend to performance monitoring's three blind spots
3. Consider relationships, perspectives, and boundaries

These general principles have much in common with the collaborating, learning, and adapting (CLA) approach being promoted as part of the Program Cycle within USAID. CLA embraces a broad range of efforts to engage with other development actors in the pursuit of development results. The complexity-aware monitoring principles often wed CLA tenets with the rigor of tested monitoring methods.

Synchronize Monitoring with the Pace of Change. As the pace of program adaptation quickens (and sometimes slows) to match the pace of change in the context, monitoring must adjust if it is to continue to provide useful information. Experience and engagement in the operating environment is the best way to gauge the pace of change. In highly dynamic contexts, monitoring may take place on a frequent, or even on-going, basis. However, some significant results may require considerable time to emerge; it makes sense to monitor for these results less often.

In both fast-paced and slowly evolving circumstances, effective program management requires indicators that provide information before or during important changes in the project and environment. In complexity, leading or coincident indicators[7] are more useful. If you rely on lagging indicators, by the time you know about a change, it may be too late to act. For example, if the results are to be policy changes, a leading indicator would be the first signs of progress in the policy process such as a key actor identifying or debating a policy issue. Leading indicators act as an early warning system to alert program managers of the need for course correction.

- **Leading indicators** provide information before the result takes place.
- **Coincident indicators** yield information at about the same time as the result.
- **Lagging indicators** provide data after the result takes place, often with considerable time lag due to data collection routines and long result chains.

Attend to Performance Monitoring's Three Blind Spots. As part of the Program Cycle, monitoring is organized primarily around answering questions about the progress of interventions towards desired results according to predetermined implementation plans. Consequently, monitoring systems tend to focus primarily on intended outcomes, intervention(s) as the dominant causal factor, and the causal pathways outlined in results and logical frameworks. Emphasis on proving a causal link between the intervention and outcomes along predicted, linear causal pathways means that performance monitoring is virtually blind to 1) a broader

7 http://en.wikipedia.org/wiki/Economic_indicator

range of outcomes associated with the intervention or system (intended, unintended, positive or negative), 2) alternative causes from other actors and factors and 3) the full range of non-linear pathways of contribution.

This narrow focus makes sense for monitoring the simple aspects of strategies and projects; however, ignoring unintended results, alternative causes, and the multiple pathways of contribution is perilous for complicated and complex aspects of projects and contexts. Some may argue that the benefits of performance monitoring as practiced in the Program Cycle outweigh the limitations posed by its three blind spots, looking to evaluation to supplement performance monitoring's narrow focus. Unfortunately, evaluation does not currently play this role in USAID. The recent meta-evaluation of 340 USAID evaluation reports found that only 15% reported on unplanned effects, and only 10% discussed causes in addition to USAID interventions that might be contributing to results,[8] Regardless, in programming environments of substantial complexity, where the ability to predict outcomes and causal pathways is low, evaluation is insufficient to steer effective implementation. Complexity-aware monitoring has a critical role in tracking a fuller range of outcomes, causal factors, and pathways of contribution.

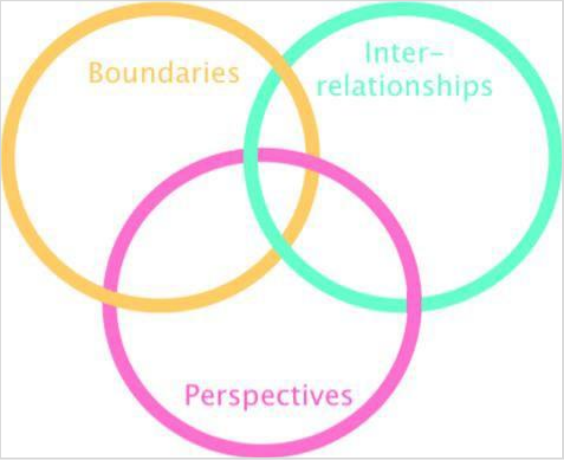

FIGURE 3: THREE CENTRAL SYSTEMS CONCEPTS
Source: Wilson-Grau, R., 2013

Attend to Relationships, Perspectives, and Boundaries.[9] Three central concepts enable us to understand the scope, focus, and intent of the systems field. Consideration of these three concepts can guide monitoring in complexity (Figure 3). The concept of relationships emphasizes that the essential features of any system lie in the dynamic interconnections among parts, not in the individual parts themselves. Does monitoring track the structures, processes, and exchanges linking actors and factors within a system? Different actors in the system have different perspectives about the relevant relationships in a system, that is, they see, describe, experience, and value those relationships differently. Does monitoring provide information on the different perspectives within a system? Different actors may also define a system differently and include different elements and relationships. Boundaries determine what is included within the system and what is considered outside the system. Does monitoring provide information that is useful for the consideration of what is in and what is outside the system?

To gain the maximum benefit from these systems concepts, they must be used synergistically. Apply the three together and to each other: What are the perspectives on the boundaries and relationships? What are the relationships between perspectives and boundaries? What are the boundaries on perspectives and relationships — whose perspectives and relationships are taken into account? Every monitoring endeavor makes choices between which actors and factors are deemed relevant and which are not, which relationships it includes and which it excludes, which perspectives are honored and which perspectives are marginalized. When designing monitoring systems, these three systems concepts can help you reflect on those choices. Systems concepts can also help when using monitoring data to guide implementation. Incorporating consideration of relationships, boundaries, and perspectives helps to disrupt any single interpretation of a situation, and provokes more creative thinking and collaborative problem-solving.

8 Hageboeck, M., Frumkin, M., & Monschein, S. (2013). Meta-Evaluation of Quality and Coverage of USAID Evaluations 2009-2012, Management Systems International under subcontract to DevTech Systems, Inc. for USAID Contract No. AID-OAA-M-11-00026.
9 This section is based on Williams, B. (2011). All methods are wrong, some methods are useful. Systems Thinker, 22(4).

For example, the small-scale producer, buyer, exporter, financier, and end-market user each experience a market system differently. It can be useful to consider these roles as each representing a distinct perspective with a unique understanding of the boundaries of the market system. The small-scale producer thinks primarily about her crop, buyers, and input services, like the shop where she purchases fertilizer. The exporter may not consider input services; instead he will emphasize the role of shipping and transport companies in the same market system. The financier will draw the market system boundaries to include debtors, creditors and banks.

Likewise, each perspective will view key relationships in the system differently. A buyer who offers a sole-source contract to a small-scale producer considers the provision of credit and harvest equipment as a fair exchange for a lower-than-market crop price. A small-scale producer may experience significant financial duress under this arrangement and consider the relationship exploitative. Another buyer views the imbalanced relationship between contract buyer and producer as a business opportunity and seeks to engage the producer in side-selling. Thus, key perspectives on a market system, and its boundaries and relationships may be taken into account when designing a project to affect change in the market system, and should be considered during project monitoring as well.

Participatory monitoring approaches are one way to put systems concepts to work in monitoring. Projects designed to achieve ambitious development objectives necessarily involve and affect a diverse array of stakeholders who bring a variety of perspectives about and relationships to each other and the project. In participatory monitoring, tasks are distributed among stakeholders to allow for variety in content, analysis, interpretation, and uses of data to achieve outcomes. Participatory monitoring may contribute to the ongoing negotiations among stakeholders needed for steering a project effectively in complexity.

PROMISING COMPLEXITY-AWARE MONITORING APPROACHES

This Discussion Note recommends five approaches to complexity-aware monitoring for USAID projects and strategies. The list is not comprehensive, but is intended as a starting point.

1. Sentinel Indicators
2. Stakeholder Feedback
3. Process Monitoring of Impacts
4. Most Significant Change
5. Outcome Harvesting

These monitoring methods are appropriate for projects or activities (both those carried out by an implementer and those undertaken by a Mission). Several of the methods may also be useful at higher levels of a country strategy results framework.

Sentinel Indicators. Sentinel indicators are the most basic way to complement a LogFrame or results framework-based performance monitoring system with a complexity-aware approach.[10] The concept of sentinel indicators is borrowed from ecology where it refers to an indicator which captures the essence of the process of change affecting a broad area of interest and which is also easily communicated.[11]

[10] Hargreaves, M. B. (2010). Evaluating system change: A planning guide. Methods Brief. Princeton, New Jersey: Mathematica Policy Research, Inc.
[11] Glossary of environmental science. (n.d.) In *Wikepedia Glossary of Environmental Science*. Retrieved from http://en.wikipedia.org/wiki/Glossary_of_environmental_science

For example, ecologists may designate a species as a sentinel of the overall health of an ecosystem.[12] Plants or lichens sensitive to heavy metals or acids in precipitation may be used as indicators of air pollution. A sentinel indicator facilitates monitoring and communicating about complex processes that are difficult to study. As a proxy, however, this type of indicator provides incomplete information, and judgments about complex processes or entire social systems based on a single indicator can be dangerous. Therefore, a sentinel indicator should be used to trigger further observation or probes.

A sentinel indicator:

- Represents processes of change that may be difficult to study in their entirety
- Is easily communicated
- Signals the need for further analysis and investigation

The identification of sentinel indicators begins with constructing a holistic picture of the project or strategy and the system that includes results, as well as causal factors and pathways not represented in the logic model's single causal pathway. One way of developing that holistic picture is to map the multiple causal pathways and feedback loops that link the project with actors and factors in the broader context. A number of approaches may be used to draft a map or picture of the intervention in relation to its context,[13] but none of them should privilege the Agency's programming interventions. Instead, the system map should include a diverse array of actors and influencing factors, with special attention paid to alternative perspectives and descriptions of how things work. At the strategic level, a system map might start with the narrative description of the development hypothesis underlying a development objective. The level of detail used for this will likely be greater than that included in the results framework. At the project level, the problem analysis may include similar information. In either case, special attention should be paid to assumptions underlying the theory of change because these capture the interactions between a project or strategy and the system in which it is embedded.

Sentinel indicators are placed at critical points in a system map to help monitor and inform the mutually influencing relationship between the program and its context. Several methods can be used to select or place sentinel indicators. Savvy development practitioners dealing with complexity often know where and when to watch for critical information that will help them steer their project or alert them to change direction. These critical points are similar to leverage points – another systems thinking concept useful for operating effectively in a complex system. According to Meadows (1999, p. 1), "leverage points are "places within a complex system (a corporation, an economy, a living body, an ecosystem) where a small shift in one thing can produce big changes in everything." Meadows cautions that there are no quick or easy formulas for finding leverage points, and that many are counterintuitive. Ongoing engagement with and study of a system is critical to identifying leverage points. Like "game-changers," sentinel indicators may not require targets and their effect on the system is not predetermined. [14]

[12] Jaffe, D., McDonough, C, ,Watzin, M., & McGinley, M. (2009). Indicator species. In C. J. Cleveland (Ed.), Encyclopedia of Earth, Washington, D.C.: Environmental Information Coalition, National Council for Science and the Environment. Rev. June 11, 2012. Retrieved from http://www.eoearth.org/article/Indicator_species?topic=58074.

[13] The Agency will be experimenting with several approaches to drafting such representations of the intervention in relation to its context, including the Rich Picture approach described in Williams and Hummelbrunner (2011, pp. 245-246).

[14] "Game-changers" is a CLA term used to refer to an event likely to have a significant effect of unknown nature on development results. The effect of a game-changer is unknown because it represents a complex aspect of the situation. The occurrence of game-changing events should trigger the re-assessment of theories of change and assumptions underlying country strategies and project designs. For example, an influx of refugees may have a significant effect on the ability of a country to achieve specific economic results, but not all effects can be predicted.

Systems thinking principles can be applied with sentinel indicators in several ways. First, sentinels signal changes in the relationships among actors and factors in a situation. Second, sentinel indicators can be chosen to represent key perspectives separate from those of USAID. Third, sentinel indicators can be useful when placed outside the boundaries defining a project or strategy. Results frameworks and LogFrames describe Agency-funded interventions as the primary causal element in a system. Where the Agency's influence is limited, complexity-aware monitoring can provide critical information related to complex interrelationships in the programming context, different perspectives, and factors outside the influence of intervention.

There is an inherent tension between indicator-based monitoring and complexity: indicators describe what we predict or expect will change, but complex aspects of a situation make it difficult or impossible to predict what will happen. Thus, the placement of sentinel indicators should be reviewed regularly and can be expected to change as the program evolves.

Stakeholder Feedback. Monitoring approaches that privilege feedback from stakeholders or make use of participatory methods are particularly valuable in complexity. Complex aspects of systems are characterized by a diversity of perspectives about desired results and pathways to achieve results. Diverse perspectives are important for at least two reasons. First, in complexity, knowledge of the system is partial and predictability is low. Second, how actors perceive a situation motivates their behavior. Understanding the system from different perspectives will help any single actor create a more holistic and useful picture.

Stakeholder feedback may involve a one-time measurement or an ongoing system. Examples of stakeholder feedback include citizen report cards, community scorecards, client surveys or other forms of collecting opinions.[15] Feedback systems might track the changes in the beneficiaries and partners that the intervention works with most directly.[16] Alternatively, feedback may target those excluded from or marginalized by the program as a means of questioning whether the boundaries of a strategy or project have been drawn in the most useful way.[17] It is particularly worthwhile to involve partners, beneficiaries, and other stakeholders in re-defining indicators or criteria of success.

Collecting stakeholder feedback can be challenging. Sampling errors may include failure to properly identify the relationship between a respondent and an intervention, or capturing the responses of dominant individuals or groups only. Obtaining feedback may be costly and logistically or technically difficult to achieve. Measurements can be misunderstood and misreported. For example, when citizens report reduced corruption, does it mean that incidents of corruption have actually declined, or that corruption has simply gone underground or shifted to new practices? Despite these challenges, the collection of stakeholder feedback is worthwhile because it provides information that is especially valuable for dealing with complexity.

Process Monitoring of Impacts (PMI) is more comprehensive than either sentinel indicators or stakeholder feedback for capturing the complexity overlooked by LogFrames and results frameworks. As its name suggests, the method focuses on monitoring results-producing processes. According to Williams and Hummelbrunner "It is essentially about identifying processes considered relevant for the achievement of results or impacts and then monitoring whether these processes are valid and actually taking place." [18] Impact-producing processes describe how a result at one level is used by specific individuals or

15Jacobs, A. (2010). Creating the missing feedback loop. IDS Bulletin 41:6. Institute of Development Studies. Oxford.
16Earl, S., Carden, F., & Smutylo, T. (2001). Outcome mapping. Ottawa: International Development Research Centre. (p. 1)
17These types of questions are suggested by a systems thinking tool called Critical Systems Heuristics. (See Williams and Hummelbrunner, 2011).
18 Williams, B., & Hummelbrunner, R. (2011). Systems concepts in action: A practitioner's toolkit. Palo Alto, CA: Stanford University Press, p. 92.

organizations to achieve results at the next level. In a sense, impact-producing processes take place between results in a LogFrame or results framework.

Like sentinel indicators and stakeholder feedback, PMI may be used to complement, rather than replace, performance monitoring systems. Theoretically, the method could be used at any level of the LogFrame or results framework. However, it seems particularly useful to outline the processes by which outputs are used by targeted beneficiaries or partners to produce the first level of results, since this is foundational to the entire project design and strategy. In this case, outputs are linked with the results they are intended to "cause" through a description of the processes by which partners or beneficiaries are expected to use the outputs. PMI is useful at the project level because the method can be used across a large number of activities and actors.

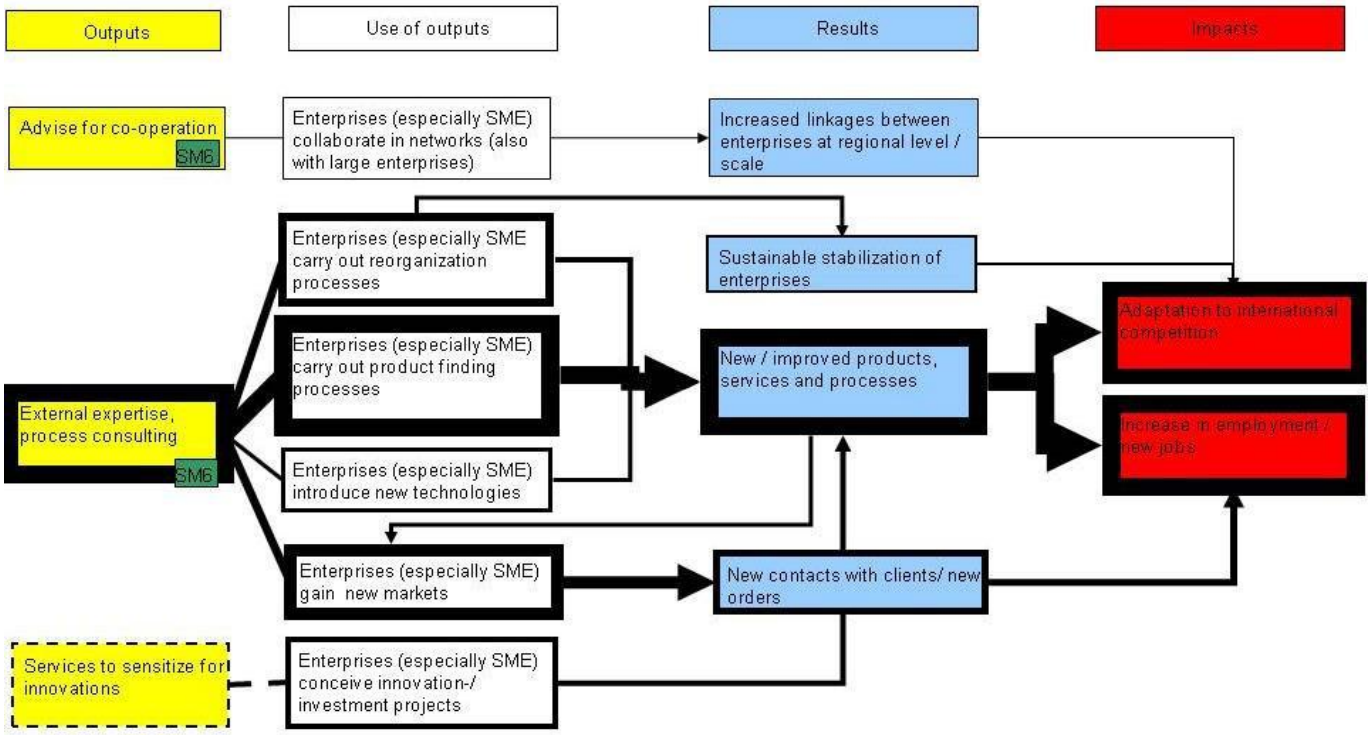

FIGURE 4: SAMPLE PMI LOGIC MODEL

Source: Williams, B., & Hummelbrunner, R. (2011). Systems concepts in action: A practitioner's toolkit. Palo Alto, CA: Stanford University Press, p. 101.

PMI involves drawing a logic model that includes outputs, first level results, and known processes that transform outputs into intended results (Figure 4). The logic model also includes any known context factors that affect the achievement of first level results, and feedback loops between the project and contextual factors. Rather than measuring a single data point (an indicator linked to a result), monitors track the processes and interrelationships represented in the logic model, which are referred to as the "area of observation." Monitoring of these areas of observation enables project managers to track results-producing processes together with the influence of contextual factors so that the project can be steered effectively long before performance monitoring data is available. Indicators may be assigned if the nature of the process under observation makes that meaningful.

In the example illustrated in Figure 4, three project outputs (in yellow) contribute to the achievement of six results on the right (in blue and red). The predicted impact-producing processes linking outputs to results are detailed in the column "use of outputs." In this example, outputs, processes and results were weighted according to the amount of budget allocated to activities associated with each output and result. Weighting is indicated by line thickness; such weighting is optional.

PMI's attention to known or predicted processes makes it suitable for complicated aspects of projects and strategies. To be useful in complexity, the method must also be attentive to emergent processes. Because it is impossible to address the three blind spots (emergent outcomes, alternative causes, and multiple, non-linear pathways of contribution) at all levels of a strategy or project simultaneously, PMI bounds the area of observation considered most critical to project success. Monitors must be attentive to both the known (complicated) and unknown (complex) results-producing processes within an area of observation.

PMI addresses several weaknesses of performance monitoring in complexity. First, PMI tracks the occurrence of impact-producing processes long before changes would be apparent in the corresponding performance indicator. Second, for USAID project designs and strategies, known results-producing processes may be outlined in assumptions or the narrative of the development hypothesis. However, they are not generally described in sufficient detail to capture the complex interactions between the project and its context. The logic model describing processes in relation to their context helps to identify alternative causes, multiple causal pathways, and feedback loops. These richer pictures help managers to modify implementation strategies as necessary even in complexity. Third, PMI recognizes that the impact-producing processes included in the logic model are incomplete estimations. The method expects that new processes may emerge and that project implementation will adapt to emphasize desirable processes and results.

By accounting for factors outside the project that influence results, and including feedback loops between those contextual factors and the project, PMI addresses two of the common blind spots of classical performance monitoring. It is, however, still focused on intended results and may not catch additional significant results, either positive or negative. PMI users should remain on the lookout for unintended results, which will likely increase with the degree of complexity. Attention to different perspectives about results and processes is one way to achieve this.

PMI applies systems thinking in a number of ways. The method is premised on a boundary setting exercise that delineates the area of observation between results. The area of observation is intended to capture complex interrelations between an intervention and its context. Various perspectives on an area of observation may be represented by monitoring data, as stakeholders report on the processes from their point of view. Monitors should also be reflective about the different perspectives on the boundary that defines the area of observation.

In a dynamic, complex system, results and causes are intertwined, emergent, and recursive. The PMI method requires extensive awareness on the part of the user to capture all results. In contrast, approaches such as **Most Significant Change** and **Outcome Harvesting** are attendant to all results, whether intended or unintended, positive or negative. These complexity-aware methods seek to discover results without reference to predetermined objectives, and work backwards to determine the contribution.

Most Significant Change (MSC) is a participatory monitoring and evaluation technique that involves the collection and analysis of stories describing the most important project outcomes. The method works well when adaptive management practices in different or dynamic contexts lead to diversity of implementation and outcomes. The method captures differences in development outcomes across sites and

time, as well as different perspectives on the same outcomes. MSC is particularly useful when different interpretations of significant change are considered valuable.

Instead of measuring indicators, the method collects and analyzes qualitative data on broadly defined "domains of change." Domains of change allow people to come to a general agreement about what to track without being too prescriptive. Domains of change point to where to look for change, but not exactly what change will look like. Project staff most directly involved with each domain of change collect stories from beneficiaries, partners, or participants. Story collectors ask questions such as the following: "During the last period, in your opinion, what was the most significant change that took place for participants in the project?" Respondents describe both the change and the reasons they consider it significant.

During analysis, different groups are assigned to review the stories and select those they consider most representative of significant change in a given domain of change. The group's selection criteria and analysis are appended to each story selected. The results of the first round of analysis are shared among all the groups participating in analysis, so that all groups come to understand the others' selection criteria. The process may involve another round of analysis to further reduce the stories and refine selection criteria. When used as an evaluation method, the process is capped by a meta-analysis of the selected stories and the selection criteria.

As originally conceived, MSC operates within hierarchical organizational structures. In this case, only the stories selected by the group(s) at the top of the hierarchy are ultimately designated as representing the most significant change. However, the technique can be adapted for more collaborative arrangements across various organizational structures, including those likely to emerge as a result of USAID's new project structure. It could be used, for example, to include key stakeholders related to the development objective who are not in a contractual relationship with the Agency. Regardless of the implementing situation, MSC supports critical thinking at the field level (closest to the domain of change), and encourages discussion and feedback among levels and functions in the organization, or among the various stakeholders collaborating to achieve a common objective.

Selected stories are verified, often by a second site visit, before final analysis is completed. When it is not feasible to verify all selected stories, a sample may be used. Stories may be quantified in several ways to provide information on the generalizability of individual stories. Quantification may take into account relevant performance monitoring indicators.

MSC is compatible with a systems thinking lens. The technique draws on the diversity of perspectives and backgrounds of those involved in a project or the context in which it is effecting change. The values and perspectives of different stakeholder groups are represented in the criteria for determining a significant change, and the analysis process makes explicit the different perspectives on a project and its results. Boundaries are considered at two stages: when defining the domains of change and during analysis. During analysis, groups may critically consider the boundaries of the domains of change as they reflect on what constitutes the most significant outcome. The meta-analysis stage generates insights about how groups operating in the system related to the project's development objective can adjust and align their relationships to achieve significant outcomes.

Outcome Harvesting. Like MSC, Outcome Harvesting is a participatory monitoring and evaluation method that enables users to identify, verify, and make sense of outcomes with or without reference to predetermined objectives. Outcome Harvesting, however, puts more emphasis than MSC on verification and on identifying and describing contribution.

Outcome Harvesting emphasizes utilization-focused[19] practice, engaging the primary user – the one who requires the findings to make decisions or take action – throughout the process. The design stage of the process focuses on defining actionable questions, the answers to which will be useful for such purposes as improving or modifying the intervention being monitored. On the basis of the selected questions, both the primary user and those conducting the Outcome Harvest agree on what information is to be collected and included in outcome descriptions. Each outcome description explains how a specific change agent contributed to changes in the behavior of particular individuals, groups, organizations or institutions, and describes what changed in these social actors' actions, relationships, policies or practices.

Data collection is an iterative process that involves reviewing secondary sources, collecting new evidence, drafting outcomes, and engaging with the key informants – the change agents. Monitors succinctly describe changes that have occurred in social actors and how the change agent contributed to these changes. In these preliminary outcome descriptions, they include questions for review and clarification by the change agent.

After collecting evidence of outcomes (positive and negative), the monitor works backward to establish a plausible cause-effect explanation of how the project or intervention contributed directly or indirectly, partially or (rarely) wholly, intentionally or unintentionally, to each change. The methods used are analogous to those used in forensics, criminal justice, epidemiology, anthropology, archaeology, and geology.[20] To ensure that the harvested information is credible enough for the intended uses, as well as to enrich it with other perspectives, the monitor validates or substantiates the outcomes with knowledgeable, independent sources.

When using Outcome Harvesting as a monitoring tool, outcomes can be collected and verified as they occur. Change agents can alert monitors when significant behavior changes become evident. A cursory analysis of substantiated outcome descriptions may be sufficient for monitoring purposes. When used in evaluation, outcome descriptions are analyzed thoroughly and interpreted through the lenses of mission, goals, or strategies and used to answer the actionable evaluation questions.

Outcome Harvesting employs systems thinking concepts. The method considers multiple perspectives about who and what has changed, when and where change has occurred, and how the change was influenced. The initial actionable questions represent the perspective of primary intended users and thus initially define what will be monitored. The perspective of the primary user is then compared with that of the change agent in the outcome description, and with the account of the substantiators. In the final stages, all three perspectives are considered in analyzing and interpreting outcomes to answer the actionable questions agreed with the primary users. Relationships between actors and factors in a system are considered when determining plausible contribution of social change agents to outcomes. The boundaries drawn to delineate an outcome and its relevant context may be considered and reflected upon.

[19] Patton, M.Q. (2012). *Essentials of utilization-focused evaluation*, Los Angeles, CA: Sage.
[20] In the field of evaluation, Outcome Harvesting represents an adaptation and development of Michael Scriven's (1991) goal-free approach and the General Elimination Method.

CONCLUSION

Five monitoring methods – the use of sentinel indicators, stakeholder feedback, process monitoring of impacts, Most Significant Change, and Outcome Harvesting – can provide data useful for steering interventions operating in complexity towards results. Premised on an understanding of social change as a complex process involving multiple and mutually influencing factors and actors, these methods generate the information necessary for both accountability and learning for complex aspects of programs and contexts.

> **Evaluation** is the systematic collection and analysis of information about the characteristics and outcomes of programs and projects as a basis for judgments to improve effectiveness, and/or inform decisions about current and future programming. The purpose of evaluations is to ensure accountability to stakeholders and learn to improve effectiveness (ADS 203.3.2).

Complexity-aware methods can be used in conjunction with performance monitoring. Performance monitoring works for simple (but not necessarily easier) aspects of strategies or projects where cause-effect relationships are known and agreement on problems and solutions is high. When USAID staff identify components of strategies and projects that do not meet these criteria, they may consider employing complexity-aware monitoring approaches.

Complexity-aware methods may differ in important ways from performance monitoring as practiced in USAID. For example, Most Significant Change and Outcome Harvesting are indicator-free approaches, while PMI and stakeholder feedback can be used with or without indicators.

Indicator-free monitoring is not completely foreign to development. Development workers often monitor the underlying assumptions of a strategy or project by tracking press reports, statements on the record by parliament members, incidents of politically motivated violence and street protests, or participation levels in markets. Indicator-free monitoring methods are often resource-light versions of recognized evaluation methods carried out with increased frequency.

Most Significant Change and Outcome Harvesting are also goal-free methods, that is, they capture outcomes without reference to predetermined results. When used in combination with a systems thinking lens, sentinel indicators, stakeholder feedback, and PMI may also point to unintended outcomes. Openness to a broader range of results is an asset of complexity-aware methods critical for those aspects of projects and contexts where predictability is low.

According to the ADS (203.3.2), the purpose of performance monitoring is to "reveal whether desired results are being achieved and whether implementation is on track." Complexity-aware monitoring can play a similar role, however, implementation is no longer about compliance with a detailed strategy or project implementation plan, and intended results may be refined or revised as implementation progresses and unexpected results emerge. Complexity-aware monitoring is ideal for the probe-sense-respond approach to implementation. But when complexity-aware monitoring is used in conjunction with adaptive management, decisions are made on an ongoing basis to guide progress towards one result over another. Typically, judgments to improve effectiveness or inform decisions about current programming fall within the sphere of evaluation; in complexity, these judgments are made iteratively. Adaptive management in complexity may blur the lines between monitoring and evaluation further, leaving us to wonder: What's the difference between complexity-aware monitoring and evaluation? We may need to revisit this question once the Agency has more experience applying complexity-aware principles and methods in the field.

WHERE TO GO FROM HERE?

This paper is intended to raise questions, stimulate dialogue, and – most of all – inspire experimentation. Most of the principles and promising approaches provided here are backed by a significant body of theory and practice, but they have not been used to monitor USAID strategies and projects. There is much to learn about whether and how they can be applied successfully in the Agency.

USAID staff wishing to experiment with one or more of the methods highlighted in this discussion paper should consult the next section for guidance on how use these new methods and approaches in a way that will contribute to learning about their effectiveness in USAID.

In the words of Arthur Ashe, "Start where you are. Use what you have. Do what you can." Let's go!

SHEPHERDING NEW METHODS

Planning to Learn. Those who champion a new method to address a measurement challenge or an information gap demonstrate the kind of leadership necessary for the Agency to be a premier development organization. These champions create and sustain the space to experiment, recruit, and support early adopters, and guide the process of learning and knowledge management. However, careful consideration of a few key points will increase the likelihood that early trials of a new method contribute to learning about its usefulness for development results.

BEFORE A FIRST TRIAL

1. **Study the method**.
    a. Talk with individuals or organizations who have used the method in the field. Get as much practical information as possible about each of the following:
        i. What purposes does the method serve and questions does it answer?
        ii. Under what conditions is the method feasible?
        iii. What resources are required?
        iv. What special challenges have users encountered?
    b. If necessary, consider hiring an expert to provide answers to these questions, keeping the level of research commensurate with your needs and budget.

2. **Develop questions about the first trials.** Using what you learned about the method in Step 1, identify questions you still have about the method's usefulness and feasibility for your program's context and development stage. Document these questions so that they can be addressed during first trials. Seek out different points of view regarding the method.

3. **Leverage resources for trials of the new method**. Fund trials of new methods and provide technical and management support. Adequate resources help to offset the risks for early adopters.

FIRST TRIAL

1. **Identify opportunities** for testing the method. Be sure to match trial opportunities to the strengths of the method. Consider if this method is suitable for addressing the monitoring information needs or the evaluation purpose and key questions. Even a promising method will perform poorly if it is not suitable for the purpose.

2. **Recruit early adopters.** Effective early adopters see promise in the method, are willing to take risks, and have influence among their colleagues. They exhibit both excitement about the trial and healthy skepticism about the method.

> *"You have to have a good sales pitch. Right up-front, people want to know about cost, sustainability, and value-add. We have to have good answers to those questions including a compelling description of what it can add."*
>
> - USAID/OTI Staff

3. **Explain the practical applications of the method**. Clearly communicate the method's (presumed) added value, and outline the costs, both financial costs and the time required to implement and manage a trial. Prepare an honest, but persuasive pitch about the method's strengths and limitations based on what is known so far. Explain the questions that remain about the method (see above) and why it is worthwhile to answer them through a trial of the method.

4. **Provide technical support to first trials.** Invest time in carefully defining the purpose and questions of evaluations considered for a new methods trial. Do your best to ensure that the method is suitable to address these needs.

5. **Give it more than one chance**. Even if the method fails in one place, consider additional trials in different settings and circumstances. Analyze why it failed in one context to help you identify conditions that might make it more effective, and then seek other contexts where those conditions are in place. Document your analysis and share with others who are interested.

LEARNING EVENTS

1. Take time to **reflect and document** your observations about managing implementation of the method. Request that those applying the method keep an implementation diary.

2. **Consult with peers** who have used the method to problem-solve during a trial. Convene staff involved in trials and those outside the agency with experience in the method to **identify lessons learned** and compare lessons from several trials.

EXPAND THE AUDIENCE

If the method is deemed suitably promising, spread the word.

1. **Be thoughtful about the first introduction** of the method to a broader audience. Choose a time and setting when you will have the full attention of your audience.

2. **Communicate clearly about the method.** Describe when is it useful, what it is likely to cost, and other practical lessons learned from trials. Share learning and experience through the Agency's various knowledge management platforms, including ProgramNet, Learning Lab, the Evaluation Interest Group, and SILK.

3. **Recruit decision makers and those with influence** when taking promising methods to a larger scale. Identify decision makers (those in charge of resources and setting policy) and potential champions (those with influence) who might consider the new method favorably. Develop a strategy to engage decision makers with the learning. Sharpen messages about why specific approaches lead to more solid evidence and better development planning.

4. **Find ways to make it easier for people to apply, adopt, and adapt these methods.** Identify the practical barriers for those trying the method for the first time. Consider drafting a tips sheet, scope of work, or request for proposal or other tools that ease the planning and management of a new method.

ACKNOWLEDGEMENTS

My gratitude is wide and deep, but space is limited. Many thanks to my colleagues in USAID's Office of Learning, Evaluation and Research for the opportunity to write this paper and their support throughout the process. I am especially indebted to Cynthia Clapp-Wincek, Tjip Walker, Melissa Patsalides, Elizabeth Callender, Travis Mayo, and Stacey Young. This paper would not be possible without the generosity of four leaders in the area of complexity-aware monitoring and evaluation -- Bob Williams, Ricardo Wilson-Grau, Richard Hummelbrunner, and Patricia Rogers. Bob and Ricardo were exceptionally steadfast as I wrestled with putting ideas on the page. Thanks to the fine folks at DevTech's Program Cycle Service Center. Jackie Greene was an unfailing source of encouragement. This paper began with a series of inspirational conversations with USAID staff who operate in complexity every day. There is nothing I would rather do than collaborate with these exemplary people.

FURTHER READING

COMPLEXITY

Westley, F., Zimmerman B., & Patton, M.Q. (2006). Getting to maybe: How the world is changed. Toronto: Random House Canada.

Kurtz, C. F., & Snowden, D. J. (2003). The new dynamics of strategy: Sense-making in a complex and complicated world. *IBM Systems Journal*, 42(3), pp 462-483.

COMPLEXITY-AWARE MONITORING PRINCIPLES & METHODS

Davies, R. and Dart, J. (2005). The 'most significant change' (MSC) technique: A guide to its use. Melbourne. Available at http://www.mande.co.uk/docs/MSCGuide.pdf

Meadows, D. (1999). Leverage points: Places to intervene in a system. Hartland, VT: Sustainability Institute. (December, 1999)

Rogers, P. (2011). Implications of complicated and complex characteristics for key tasks in evaluation. In K. Forss, M. Marra, and R. Schwartz (Eds.), *Evaluating the complex: attribution, contribution, and beyond*. New Brunswick, New Jersey: Transaction Publishers.

Scriven, M. (1991). Prose and cons about goal-free evaluation. *American Journal of Evaluation* 1991 12(55). Available at http://aje.sagepub.com/content/12/1/55.

Williams, B., & Hummelbrunner, R. (2011). Systems concepts in action: A practitioner's toolkit. Palo Alto, CA: Stanford University Press.

Williams, B. (2011). All methods are wrong, some methods are useful. *Systems Thinker*. 22(4).

Wilson-Grau, R., & Britt, H. (2012). Outcome harvesting. Cairo, Egypt: The Ford Foundation's Middle East and North Africa Office. Available at http://www.outcomemapping.ca/resource/resource.php?id=374.