Organizational Capacity Development Measurement

Executive Summary

This document provides a set of recommendations for measuring the results of USAID-supported organizational capacity development efforts. It provides grounding principles and background that inform its recommendations. The purpose behind these recommendations is to improve the consistency with which USAID program managers and partners appropriately measure organizational capacity development, enabling more effective learning from and accountability of capacity development programming across the Agency.

The recommendations cover aspects of both what to measure and, to an extent, how to measure it, but leave large space for staff to interpret and apply them as appropriate for the particulars of their programming. The recommendations describe an approach rather than a single indicator as most appropriate to measuring capacity as a multifaceted topic. These recommendations offer an important step forward in thinking about why and how to invest in improving organizational performance, and in capturing the value that capacity development is adding to development.

The recommendations are:
> ➢ In defining measures for organizational strengthening, performance measures are the most appropriate area of emphasis – generally with performance expectations set jointly with the assisted organization(s). Measurement should be centered on organizational performance.
> ➢ Performance should be measured across multiple domains, including adaptive functions, to reflect capacity development investments in both short-term and long-term aspects of performance.
> ➢ An organization's performance depends on its fit in a wider local system of actors, and its interrelationships with them. Therefore, we must measure at both organizational and local system levels in order to capture the value of performance change.
> ➢ Organizational performance change is pursued in order to affect wider, systemic changes. However, attribution for change is unlikely to be provable. We should trace the credible contribution from organizational to system change with rigor.
> ➢ Some ways in which organizational capacity development will affect future performance cannot be anticipated at the start. Therefore attend to multiple pathways of change and to the unpredicted in order to perceive the full spectrum of results.

Two of these recommendations – to emphasize organizational performance as the metric for success of organizational capacity development investments, and to measure at multiple levels including organization and local system – are echoed as requirements in Agency policy guidance for monitoring.

Consensus: Capacity

What is Meant by Capacity?

USAID has no single definition of capacity, and deliberately chose not to create one during this process, for two main reasons. First, this document identifies several fundamental aspects of capacity that should inform its measurement. These fundamental characteristics and their implications are more salient to the recommendations made herein than a specific definition. Second, there are a number of excellent definitions available and in broad use which we think serve as better common reference points than a brand-new definition – most pertinently the "Five Capabilities" stemming from a major study by the European Center for Development Policy and Management, and the definition used by the book *Capacity Development in Practice*, as well as commonly-cited definitions by the UNDP and OECD:

> *ECDPM's Five Capabilities: "To achieve its development goals, every organization/system must have five core capabilities: to act and commit; to deliver on development objectives; to relate to external stakeholders; to adapt and self-renew; and to achieve coherence."*

> *Capacity Development in Practice: "Capacity is the ability of a human system to perform, sustain itself, and self-renew."*

> *UNDP Definition: "The process through which individuals, organizations and societies obtain, strengthen and maintain the capabilities to set and achieve their own development objectives over time."*

> *OECD Definition: "Capacity is the ability of people, organizations and society as a whole to manage their affairs successfully. Capacity development is the process whereby people, organizations and society as a whole unleash, strengthen, create, adapt and maintain capacity over time."*

Principles of Capacity

- *Capacity, at organizational level, cannot be understood without reference to the wider system that surrounds any organization*

Capacity as a concept can only have meaning if it describes the capacity of an organization to perform within its context – the system of other actors that an organization affects and is affected by in carrying out whatever actions it performs. Normative statements of how "organizations of type x *should* operate" must be grounded in a rich picture of the actual situation in order to support capacity development that maximizes value-added. Capacity development approaches should always reference a relevant local system as it informs the organization's current role, and

Note: Different Levels of Capacity

These recommendations center on measurement of organizational capacity. Capacity exists at several different levels – individual, organization, network, system, etc. Any organizational capacity must encompass the people within an organization and must be oriented within the local systems in which an organization is embedded. There are ramifications around measurement at other levels that can be inferred from this document, but it does not speak to other levels of capacity directly.

describe how capacity development investments aspire to create change sufficient to affect that system. Efforts to strengthen the capacity of an organization must derive from a clear understanding of the roles it currently plays within its wider context.

- *Capacity involves complexity*

Key aspects of capacity are emergent properties of how people interact within and across organizations – capacity is produced in constant and ever-evolving ways. Capacity will emerge in non-linear fashion, and at least some aspects of capacity are always in flux as the organization and its context shift. Capacity at any point in time is therefore a snapshot of a dynamic reality. In addition, capacity will be understood differently by different people within an organization, and multiple perspectives are legitimate in terms of understanding and describing what the capacity of an organization is. Because organizations themselves have complex features and are nested in complex local systems, contribution of outside capacity development efforts to performance change is the strongest claim an outsider can make – one cannot attribute changes in a complex organization or system solely to one effort.

- *Interrelationships are central to performance*

How an organization interrelates with other stakeholders is critical to how that organization performs. Any true understanding of organizational capacity must emphasize the quality and breadth of relationships that an organization and its personnel have with others and how those inform the organization's work.

- *Capacity development should be designed to link achieve performance improvement and yield change across a local system*

Note: Public Sector Organizations

With respect to public sector organizations, there is an additional consideration – the fact that any investment in strengthening a public sector organization is also, by definition, an investment in improving systemic outcomes related to public financial management and public accountability. In order to consider the relevant systemic outcomes when working with public sector organizations, therefore, it is generally useful to track the work against one or more of the systems indicators around public financial management or public accountability – most of the generally updated ones are part of the set of regularly collected Indicators of the Strength of Public Management Systems (ISPMS) coordinated through the World Bank.

Capacity development should clearly lay out a theory of change that shows how investments are predicted to contribute to performance improvement, updating that theory of change as required over the course of implementation (see ADS reference on Human and Institutional Capacity Development for more on such an approach). The theory of change should describe how investments should lead to performance improvement of cohorts, networks, markets, or relevant wider systems – including those not directly partnered with USAID – reflecting the contribution of our work.

- *Capacity development should invest in adaptive functions that help an organization thrive over time*

Because nobody can anticipate all possible future situations, a portion of any capacity development should invest in adaptive functions of an organization so that it can better meet unknown future challenges and

opportunities. An organization's drive to achieve continuous improvement, adapting to new learning and a changing context, is an important factor in its success.

- *Local ownership underpins changes in capacity*

Staff and stakeholders within an organization understand its capacity and can act to support or obstruct changes in ways that outsiders cannot. Their perspectives are therefore not only integral to perceiving and measuring the current capacity of any organization, but are essential to grounding any plans to support capacity development or organizational change.

- *Capacity development takes time*

Capacity development interventions can be conducted in relatively short timescales; however, for organizations to embed changes in ways that improve their performance takes time. This time lag should be accounted for in distinguishing between the timespan in which activities to support capacity development are conducted, and the timespan in which shifts in organizational performance are expected to become visible.

Measurement Recommendations

Taken in total, the principles of capacity and capacity development outlined above make it clear that capacity is best measured through an approach encompassing several methods rather than by a single indicator. However, it is also clear that such a measurement approach can be applied to diverse capacity development efforts and different types of organizations. Through identifying and applying a similar measurement approach, it is expected that more data can be generated to enable better monitoring, evaluation, and learning around organizational capacity development across sectors and organization types.

First, because capacity is expressed through performance, capacity development ***measurement must be centered on organizational performance.*** When USAID monitors the results of capacity development, it should monitor performance change as the appropriate metric for validating whether capacity has changed in ways that are significant. This has the further benefit of aligning incentives between what USAID monitors and what organizations aim to achieve through the development of their own capacity – neither USAID nor partner organizations seek to develop capacity for its own sake, but rather to better empower their organization to achieve its goals and objectives.

This emphasis on organizational performance does not imply that underlying process and input measures are not useful. In public sector strengthening, for example, there is great consensus around certain internal processes as being valuable in and of themselves – transparency and stakeholder engagement in planning and budgeting, matching budget execution to budget formulation, establishment and use of accountability channels, degrees of bureaucratic autonomy, and availability of appropriate inputs. However, the performance of any organization remains the most important aspect for measurement as it relates to organizational change.

Recommended tools for this include composite indicators or selected key performance indicators. Any such indicators should reflect the organization's buy-in, and ideally come from metrics it already uses to gauge its performance.

Second, performance measurement must be defined holistically, ***encompassing both the organization's performance in achieving targeted results and the organization's performance in learning, adapting, and sustaining itself over time***. An organization's performance matters in at least two senses – an organization's performance in achieving results, and an organization's performance in adapting and renewing itself in response to its changing context.

In order to identify a common language for these different dimensions of performance, the Local Solutions working group is recommending adoption of the IDRC/Universalia Framework for organizational performance that is operationalized in the Pact Organizational Performance Index (OPI). The OPI's Framework is shown here for reference, with its four domains of effectiveness, efficiency, relevance, and sustainability. Other offices and units are employing other index indicators or tools. Regardless of the tools or indicators used, ensuring a focus on performance and attention to performance areas such as relevance and sustainability that matter more over time will enable more effective monitoring.

Domains of Performance

Adapted from PACT
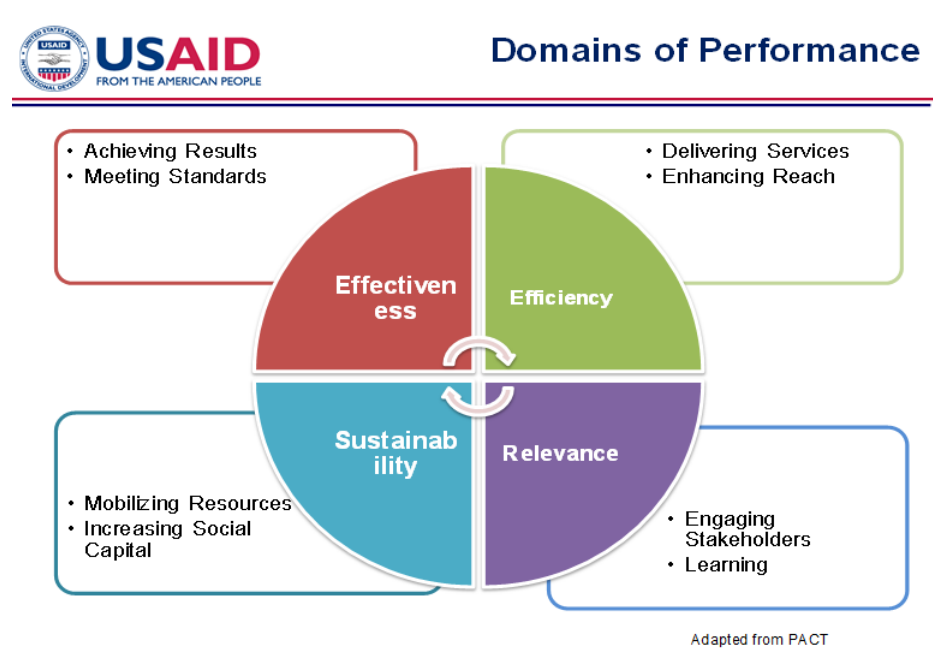
Third, the ***measurement of organizational performance must be complemented by measures of the wider local system that co-produces the development results of interest***. For the given organization, its performance horizons are shaped by the local system around it, and performance measurement depend son observing how it functions within that wider system. And to speak to the value that a given organization's performance improvement may have, one must observe how that role as well as the wider system is changing as a result of capacity development supported by USAID.

Any targets of expected performance change should derive from the activity's articulated theory of change for how organizational performance improvement is predicted to affect a wider local system. This requires a clear description of the roles in local systems that given local organizations are playing as a baseline. Further, targets for performance change (and the theory

of change relating the organization's performance to a relevant local system) should be validated with the partner organization and consensus established around targets.

For example, if USAID is supporting improved performance by public organizations providing agricultural extension services, USAID would want to measure both the performance change of those organizations and the performance of the agricultural value chains that those organizations' efforts were intended to improve. Or if USAID is supporting improved performance in budget formation and execution by selected municipalities, we would also want to measure a systems outcome such as the perceived fairness and legitimacy of the state by citizens in the target regions, or improved cost efficiency in service delivery for publicly-funded services in the target regions.

Due to the importance of interrelationships as structuring the way in which capacity emerges, it is recommended to include at least one measurement at systems level of the interrelationships between actors and how those are changing over time. Measurement of interrelationships can be either qualitative or quantitative, and may not be easy to link with targets, but relationships within the relevant system often serve as a key context indicator to be regularly reviewed and used to inform programming. Some projects have successfully used social network mapping or related techniques to visualize and quantify this type of data, and this seems a practice with high potential to add value to Mission learning. Other tools to measure systems can include wide stakeholder feedback through collection of narratives or polling data; visualization of systems dynamics or constituent parts; or indicators of system stocks and flows.

> ## Core Recommendations
>
> 1. **Measurement must be centered on organizational performance**
> 2. **Measurement performance across both achieving targeted results and in learning, adapting, and sustaining itself over time.**
> 3. **Measurement of organizational performance must be complemented by measures of the wider local system that co-produces the development results of interest.**
> 4. **The credible contribution of organizational performance change to local system change will fit a contribution paradigm.**
> 5. **Measurement approach should incorporate at least one method of perceiving unpredicted changes in performance and of validating the pathway of predicted changes.**

Fourth, *the effect of organizational performance change on local system change will fit a contribution paradigm*. Given the complexity of local systems, statements about the linkages from performance change to effect on local systems will necessarily be contribution rather than attribution. USAID can increase the rigor with which confidence is established in the contribution of performance improvement to system change through the use of multiple methods to connect organizational performance and systems change, and through gathering different perspectives on change.

Fifth, the ***measurement approach should incorporate at least one method of perceiving unpredicted changes in performance and of validating the pathway of change where predicted changes in performance occur***. This often requires deductive approaches that trace processes after change has happened. Employing these approaches also adds rigor to assertions of contribution along predicted lines. Because capacity development is an engagement with complexity, initial theories of change should be updated through the validation of pathways of change.

Even when performance change is measured where one has supported capacity development, one must gather input to validate that outside support contributed to that performance change. This entails some process tracing or other ways of looking backwards at how capacity development support was understood to yield performance change, including multiple perspectives on the same question. And since some performance change is likely in areas where it was not predicted, efforts to understand an outside contribution to performance change should include effort to examine the pathways through which change happened and to look at contributions to unpredicted performance change.

Several examples of these types of tools are captured in the Discussion Note on Complexity-Aware Monitoring, and all three blind spots of performance monitoring noted in the Discussion Note are relevant to capacity development.

Scope and Use of This Document

The approach described in this document covers organizational capacity and principles to apply when measuring its change. It is closely related to efforts to measure wider changes (across a relevant local system) and issues of ownership and sustainability to which organizational capacity can contribute, however it does not address those issues directly. Measurement following these recommendations is intended to serve as one part in a chain, and to offer more rigor for speaking to the contribution that Agency efforts to strengthen organizational capacity are making to higher level, wider system results.

This approach is applicable to any type of organization: public or non-public, for-profit or not-for-profit, formally or informally defined, of any size. Each of those factors may introduce considerations that inform the specifics of monitoring or evaluation, such as issues around data availability or time and expense of data gathering. Certain types of organization may have specific constraints that affect how their capacity is shaped and expressed. Every organization's capacity is also shaped significantly by the wider systems in which it is embedded.

It is important to emphasize that this measurement approach is informed by scholarship and practice related to capacity development in diverse sectors and organization types, and reflects the commonalities and consensus areas across those realms. It pushes practitioners to move from older mental models of capacity development that articulate best practice attributes of organizations toward an approach rooted in context and best fit, in keeping with the latest thinking in the discipline.

This document covers findings on capacity and capacity development in some detail, and as such, it is also relevant to project design and activity design. However, the focus of this document is around how measurement is done and that purpose guides the form and content. It embeds several key principles around capacity and capacity development and recommendations related to them, but the emphasis of the recommendations is on measurement.

In order to use these recommendations, Mission staff will have to identify appropriate methods to follow them. In general, these will require 2-3 related indicators or other monitoring tools, some examples of which are suggested as illustrations, which can help to capture the different dimensions outlined by the principles. Putting together such a monitoring package (or requiring such of an implementing partner) should not be significantly more expensive or time-intensive than for other programming. The regular use of the measurement data to inform programmatic adaptation will be critical.

The recommended measurement approach is illustrated by three examples included as annexes: two illustrative Project M&E Plans, and sample language for a solicitation asking for this approach to be used in applicants' proposals.

As part of a regular review with partners of their perspectives on organizational change, USAID should collect information regarding the changes linked to our support for organizational capacity development that are asserted to have occurred (and evidence in support of those changes, as appropriate), and a method can be identified as a feature of the learning plan for the project even absent a specific tool or technique. Such information will often arise naturally from a collaborative learning discussion at a portfolio review, for example, if a CLA approach is being employed in the project. It is essential that this information is valued, intentionally gathered, and documented to enable learning and accountability for the full array of effects of the USAID programming, not just those hypothesized at the project's or activity's outset. Importantly, if the data gathered is inconsistent with the initial theory of change relating the organization's performance to the wider system, this should create an opportunity to adjust the theory of change to better match the data being tracked and change processes being monitored.

Other Purposes and Their (Distinct) Tools

There are other purposes for which an external actor like USAID might support a review of some aspects of an organization's capacity. However, these purposes are distinct from measurement of capacity change, and the tools for them are also distinct. One of the consistent messages received during consultation with practitioners has been the importance of having measurement tools for capacity and capacity development that are clearly distinguished from tools for other purposes. USAID uses specific tools for purposes of risk assessment and catalyzing capacity development, some of which are highlighted below; these should be kept separate from measurement tools.

Risk Assessment and Mitigation

Risk assessment and mitigation often entails reviews of aspects of an organization's capacity and/or function, to identify how that organization functions that create risk around how they are engaged with USAID work. For example, prior to providing funds directly to a partner

government organization, USAID conducts a multi-stage **Public Financial Management Risk Assessment Framework (PFMRAF)** review that considers aspects of the organization's financial management systems, democratic accountability considerations, and looks at both the wider context in its first stage and at the specific entity's internal functions in its second stage. Prior to making an award to an organization, there is often a pre-award survey that similarly looks at a few touchstone points of potential fiduciary or programmatic risk, including internal controls, segregation of duties, accounts management, etc. The key distinction of this purpose is that it is an external review oriented around determining the level of risk entailed in partnering with the organization and approaches that can mitigate that risk. While there is a relationship between an organization improving its capacity and an organization becoming less risky as a potential partner, it is inaccurate to say that reductions in risk and improvements in organizational capacity are the same. Because risk is filtered through the lens of the USAID relationship – it is about risks in partnering with an organization for a specific, short-term purpose – it is not appropriate to substitute a risk assessment for a measurement of holistic organizational capacity or its expression.

Catalyzing Capacity Development

As part of external support for capacity development, outside actors often support "assessments" of an organization's capacity that serve as inputs to catalyze change. For example, USAID has standardized an **Organizational Capacity Assessment (OCA)** tool, and various implementers have four- or five-stage maturity models across different areas of organizational function. There are also wider industry standards, such as those supported by the International Organization for Standardization (ISO). These are all incorporated into capacity development as efforts to identify and prioritize changes to be pursued by the organization. The key distinction of this purpose is that it is less interested in capturing the ways in which organizational capacity is expressed than in motivating achievable actions to be undertaken. Most such tools are willing to sacrifice data validity for motivation. For example, the OCA tool is explicitly a self-assessment, often with limited validity, because research has found that an organization conducting a self-assessment is more likely to follow through on identified actions than an organization receiving an expert review of its capacity, even if the latter is more accurate. As a self-assessment, OCA is by definition not valid as a measurement tool, as it is expected that organizations may mischaracterize their own capacity.

In addition, because the purpose of these tools is to foster organizational change, they emphasize internal functions of organizations that are more likely to be within the manageable interest of those organizations to change. Many of the tools used in a typical HICD engagement to catalyze capacity development share features with the OCA – they are perception-based reviews through a self-assessment, and their purpose is to support and motivate change.[1] This is not to suggest that self-assessments are unimportant – quite the contrary, for an organization to perform well on critical features such as learning and adaptation, it must find ways to regularly self-assess and seek continuous improvement. However, the value in these assessments is in the actions they

---

[1] An HICD or other capacity development process may build consensus around appropriate indicators to measure organizational change and performance, which are valid for measurement purposes; these jointly-agreed indicators are kept distinct from self-assessment processes in order that they can serve as objective data sources. For more information on HICD, see ADS 201 Mandatory Reference.

motivate, not in the subjective scores or ratings they provide. Attempting to use the same tool to support capacity development and to measure the effect of capacity development introduces a tension into the tool that limits its effectiveness for both purposes. Therefore, it is not appropriate to substitute a capacity development tool for a measurement of organizational capacity or its expression.

Leveraging the Learning

How This Approach Compares to Current Practice

Presently, many Agency units support organizational capacity development, and most of them incorporate portions of the guidance within this approach. Tools and monitoring methods have evolved in recent years, as has the wider policy environment, and these allow a robust measurement approach to capture improvements in performance that is often now unrecognized.

It is worth noting at the outset that part of the rationale behind creating a common measurement approach is in order to better align Agency incentives – what USAID measures in its programming is, by virtue of being measured and made visible to project and activity managers, often what we and our partners perceive as valued - "what counts, matters." It is therefore most useful to highlight where the recommended measurement approach differs from typical Agency practice.

First, much of the CD measurement that currently occurs places emphasis on measuring capacity qua capacity rather than measuring performance change, or mixing the two together. Often there is an imported "best practice" normative model for how an organization should perform that is not relevant to the fit between a given organization and its local system. Sometimes the same tool is used to assess risks or to catalyze capacity development as well as to measure capacity change. In either case, these introduce perverse incentives into the capacity development, biasing capacity development toward compliance checklists and allowing for organizations to "signal" capacity change without truly improving performance.

Second, when performance is measured, the emphasis is often on achieving results without due attention to performance in learning, adapting, and self-renewal. This creates incentives that privilege shorter-term accomplishments and undervalue investments in sustainability. The emphasis on short-term results, and on compliance as opposed to long-term performance, has in some cases been exacerbated by recent emphasis on aspirational targets for local awards spurred by USAID Forward's Implementation and Procurement Reform (IPR). The focus on longer-term performance and connections from organizational performance to local systems change is consistent with the shift from IPR to Local Solutions already underway.

Third, in many instances, even where capacity development is pursued, Agency activity and project managers do not measure at both the organizational performance and systems outcome levels. This obscures the logic underlying the capacity development activity and makes it difficult to adjust programming when inputs are not producing predicted outputs and outcomes. This is because absent a clear theory of change around how each level was expected to affect the next, adaptation is much more difficult. For example, if the only measure of capacity

development investments in a set of hospitals is their number of patients seen after TA provision, and target numbers of patients are not reached, it is difficult to adjust absent metrics around how internal hospital improvements were intended to allow them to see more patients (and why seeing more patients is an appropriate performance measure, given the role of the hospitals in their local system and context).

Fourth, it is not yet a common Agency practice to attend to unpredicted changes or to examine the pathways of change that occurred as predicted, as part of either routine monitoring or periodic evaluation. As many important outcomes from capacity change are not predicted in advance, this reduces the perceived effectiveness of capacity development by failing to fully tell the story of what capacity development efforts have achieved. And by not validating the pathways of change that were predicted, USAID Officers miss opportunities to update their theories of change to better reflect the context.

Finally, even where USAID support for organizational capacity development otherwise follows these recommendations, the lack of any common performance indicators makes it difficult to aggregate data or identify patterns at a level beyond the individual activity or project around what capacity development support is yielding what sort of performance change, and what performance improvements are yielding changes of significance in development results.

Uses Within Projects

For any given activity, USAID project and activity designers should first have surfaced our theory of change around how USAID expects capacity development to yield performance improvement. During implementation, staff should review the monitoring data to constantly verify or update that theory of change based on what is actually happening. Clearly identifying how the results monitored cause USAID to update its theory of change – and putting more emphasis on an evolving theory of change (and related implementation approach) than fidelity to the initial theory of change – will greatly facilitate adaptive management of capacity development programming.

Where measurement of organizational performance change is carried out appropriately, again in line with the theory of change laid out in the project design and as updated through implementation, USAID will be able to relate organizational change to measurement at systems level, and thereby speak with more clarity and rigor about our contributions to achieving and sustaining ultimate results of interest.

Uses Across USAID

USAID will also be able to apply learning across the discipline of organizational capacity development more broadly – a potential area of great learning whose utility has been undervalued due to differences that have obscured key commonalities across organizational capacity development in different organization types, sectors, and country contexts. Use of one or more shared tools to measure changes in organizational performance is expected to generate much more data from which to identify patterns – even though any such shared tools would be complemented by additional indicators or tools that address particular performance changes

specific to an organization and its context. Having a common language to describe different areas of performance improvement, and a common measurement approach underlying the appreciation of the principles of capacity and capacity development, will enable greater clarity in conversations around what is working, and feed into learning at scale around capacity development.

Annex A: Selected Annotated Bibliography
Annex B: Background and Process to This Document
Annex C: Two Example Project M&E Plans Using This Approach
Annex D: Example Solicitation Language for Activity M&E Plan that Uses This Approach