

Metrics for the Haiti Stabilization Initiative

BY DAVID C. BECKER AND ROBERT GROSSMAN-VERMAAS

Stabilization in postconflict or low-conflict situations is a growing business around the world. For the United States, stabilization efforts at the moment may seem to focus on U.S. military involvement in Iraq and Afghanistan, but the recently released Quadrennial Diplomacy and Development Review noted that there are 36 active conflicts and 55 fragile states in the world. In reality, the United States supports stabilization efforts from Colombia to Lebanon through a variety of programs. Using a parallel non-U.S.-centric indicator, the United Nations (UN) now supports more than 14,000 police in 17 different countries to provide police advice, law enforcement training, and a public security presence in situations where the UN has a mandate to support a government or encourage peace-building efforts.

As more is given *for* stabilization missions, more is demanded *from* stabilization missions. With the money comes responsibility to monitor and evaluate the funds and time spent. This is not just to avoid waste and fraud, but to prove that the overall investment was worthwhile and made a positive difference. Objective and accurate evaluation provides a basis to learn from experience and decide what should or should not be funded in the future. A rigorous metrics and evaluation effort should yield evidence of progress toward accomplishing project/program goals. Without evidence, there exists no rational basis for drawing any conclusions and basing future policy or program decisions.

In a recent Government Accountability Office (GAO) review of Department of Defense (DOD)-funded stabilization programs in 28 countries, the GAO recommended that “the Secretary of Defense, in consultation with the Secretary of State and Administrator [of the U.S. Agency for International Development (USAID)], develop and implement specific plans to monitor,

David C. Becker is a Senior Research Fellow in the Center for Complex Operations, Institute for National Strategic Studies, at the National Defense University. Robert Grossman-Vermaas is Team Lead of the Social Science Group at Logos Technologies in Arlington, Virginia.

evaluate and report on their outcomes and their impact on US strategic objectives to determine whether continued funding for these projects is appropriate.”¹

Even as there is growing interest in understanding what works in stabilization or peace-building evaluation, there is growing frustration. Several things make evaluation a hard sell:

Haiti has seen six UN interventions in the last 20 years, including the use of U.S. forces on three occasions

- ❖ People do not like to measure themselves.
- ❖ No one agrees on what to measure: “stabilization,” “conflict response,” “peace-building,” or “counterinsurgency.”
- ❖ Often programs have no clear hypotheses to measure.
- ❖ There is still a great deal of confusion about the types of monitoring and evaluation.
- ❖ There is often a concern about spending limited program dollars on something that does nothing to improve results.
- ❖ Speed kills evaluations.

Despite all these problems, there are some interesting examples of metrics in stabilization that might shed some light on what works and what does not.

The Cité Soleil Case

Haiti is one of the lesser known cases where both the United Nations and the United States are involved in stabilization efforts. Haiti has seen six UN interventions in the last 20 years,

including the use of U.S. forces on three occasions. One estimate suggests that DOD has spent more than \$1 billion intervening in and occupying Haiti on different missions. It remains a fragile state by anyone’s calculation, even with a UN force of 11,000 stationed in the country since 2004. Haiti’s weak institutions and proximity to the United States exacerbate issues of drug-trafficking, mass migration, organized crime, political manipulation, and gang violence. By 2007, one particular zone, Cité Soleil, served as a critical focal point of instability, violence, and civil unrest severe enough that it was threatening the stability of the national government, but for the presence of UN forces.

Cité Soleil is a densely populated shantytown located in Port-au-Prince. The capital’s most notorious slum is regarded as one of the Caribbean’s poorest, roughest, and most dangerous areas. It is a no-go area for anyone but gang members and a kind of lawless state within a state. There are few police, no sewers, few stores, and little or no electricity. The crime, unsanitary conditions, lack of essential services, and violence that characterize this slum have become a microcosm of Haiti’s endemic problems. The majority of the estimated 300,000 of the residents are children or young adults. Few live past the age of 50; they die from various diseases, including HIV/AIDS, or of violence. The UN Secretary General has described the human rights situation in Haiti as “catastrophic.”

The Haiti Stabilization Initiative

In response to this growing political/criminal crisis, DOD, using its new Section 1207 authority of the 2006 National Defense Authorization Act, provided \$20 million from its operation account to the Department of State. The mission was to try a “new approach

to reconstruction and stabilization in Haiti by modifying the way the [U.S. Government] combines all tools at the Embassy's disposal with the goal of markedly improving security, local government capacity, and economic opportunity in Cité Soleil."

The Haiti Stabilization Initiative (HSI) was designed by an interagency team with assistance from the State Department's Office of the Coordinator for Reconstruction and Stabilization (S/CRS) with the goal of improving stability, security, the economy, and local essential services capacity in the most volatile area of Port-au-Prince. By defusing the most urgent drivers of conflict and concurrently increasing institutional capacity and performance, the government of Haiti hoped to buy time in Cité Soleil and to build the psychological and political support that it desperately required. A follow-on effect would be a more conducive environment for U.S. and international economic and social programs in the community to expand their operating environment. The endstate was to "open the doors" of Cité Soleil so that others (and the government) could run the same assistance programs that were offered elsewhere in the country, with no more risk and difficulty than anywhere else in Haiti.

According to the Proposal for the Use of Section 1207 Funding, Haiti Stabilization Initiative, there were a series of direct results anticipated in this \$20 million experiment line:

- ❖ HSI would integrate an expedited police training and professionalization program with a community-focused effort to improve governance, infrastructure, economic outlook, and law enforcement.
- ❖ Building on the U.S. Conflict Transformation Plan for Haiti, HSI

would support a broader stabilization effort aimed at shaping Cité Soleil by creating jobs, building local leadership, and developing programs for sustainable employment.

- ❖ Local governance would be strengthened by providing the means for civil servants and elected officials to provide regular basic services.

In summary, HSI was proposed as an urgent 2-year program intended to open the way for sustained and effective U.S. and donor-funded programs to operate unhindered in Cité Soleil, thereby creating a viable, stable environment.

Developing a Monitoring and Evaluation System for HSI

As the first DOD-funded 1207 project and as the S/CRS prototype effort, the Haiti Stabilization Initiative was carefully monitored to determine the successful achievement of its desired outcomes. Was it possible to do a successful civilian-led stabilization? Was it possible to do it with only \$20 million? If so, how would it be proven that it was the HSI program that made the difference? The HSI interagency team needed a measuring stick to evaluate the program. The original budget for the HSI project included funds for a quarterly survey of population, but it seemed obvious that a survey would not get an in-depth analysis of progress.

M&E for HSI: Innovation and Adaptation

One option for a monitoring and evaluation (M&E) effort was to use the Measuring Progress in Conflict Environments (MPICE) system. Luckily, this system had been developed to the point in 2007–2008 where it needed a site to test the prototype system. The system

also met two key criteria: it had to be well researched, and it had to be independent of the stabilization program's management.

The MPICE system includes a framework, collection processes, and analytical tools. A variant of the MPICE prototype system was used for the HSI M&E program. This variant was codeveloped by Logos Technologies, first under contract to the U.S. Army Corps of Engineers and then to HSI.

The MPICE framework is structured around determining conflict drivers and state/society institutional capacity, as conceptualized by the United States Institute of Peace, Fund for Peace, U.S. Army Peacekeeping and Stability Operations Institute (PKSOI), and others. The framework was introduced to the stability operations community during the Eisenhower Security Conference in 2006 and funded by the PKSOI. It was then systematically developed over 18 months with input from broad representation across the stability and reconstruction community, including State Department, USAID, DOD, United States Institute of Peace, and international partners. The premise states that if conflict stabilization and societal reconstruction is a process continuum spread between violent

(see figure). Regardless of precise terminology, the MPICE framework is intended to provide M&E teams with a capability to generate substantial insight into conflict environments and gauge progress with respect to this continuum.

To maximize its utility to many existing planning structures, MPICE divides into five traditional sectors:

- ❖ political moderation and stable governance
- ❖ safe and secure environment
- ❖ rule of law
- ❖ sustainable economy
- ❖ social well-being.

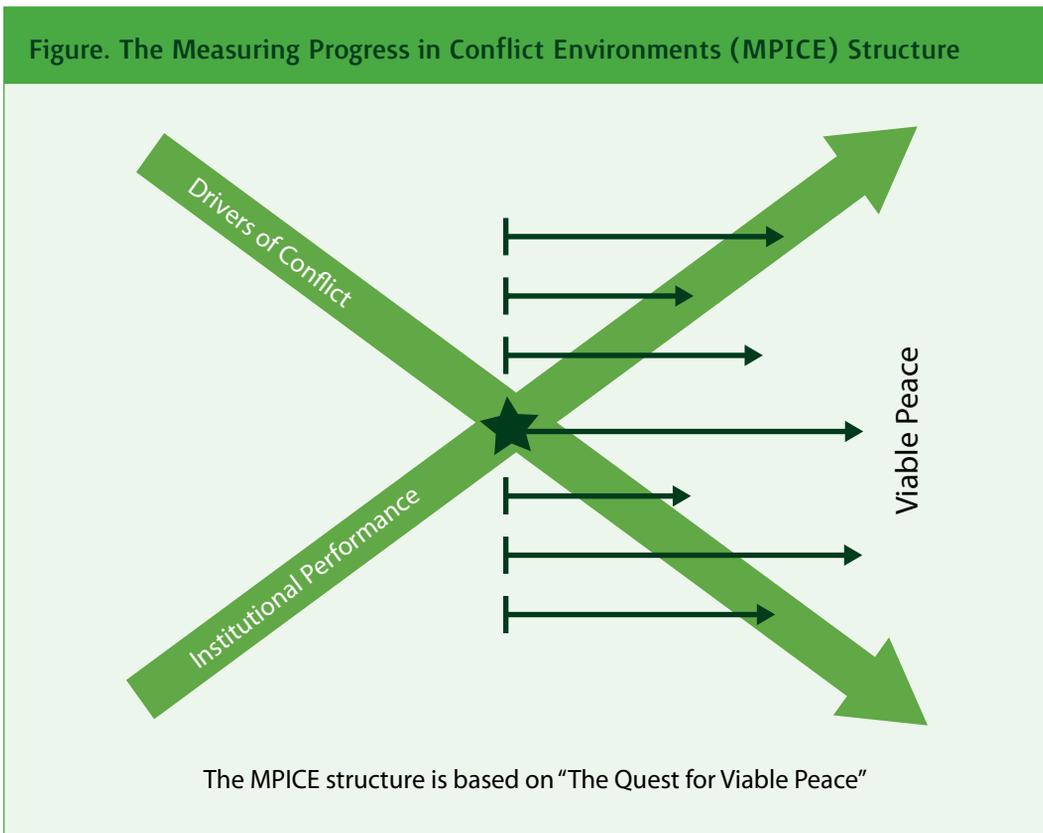
Each of these sectors divides into the two subsectors (conflict drivers and institutional performance), which flow down a hierarchy, with measures aggregating to provide indicators of progress toward the achievement of goals over time.

MPICE outcome trends are illustrated using a process in which measures are tailored to the specific stabilization environment of interest, and information is then gathered by means of several data collection methodologies. These methodologies include content analysis, expert knowledge, quantitative data, and surveys/polling data. Each of these collection methods has inherent strengths and weaknesses.

Additional methodologies can be applied depending on the environment. For example, to better assess local stakeholder perceptions of progress in Cité Soleil, Logos Technologies employed a focus group methodology to draw out coded qualitative responses to questions aligned to specific MPICE metrics, which in turn were aligned to HSI's goals. They also developed a richer, more operational version of the expert knowledge or expert elicitation methodology.

data collection methodologies include content analysis, expert knowledge, quantitative data, and surveys/polling data

conflict and sustainable security at opposite ends, viable peace should be considered the middle or “tipping point” where external intervention forces can begin to hand over driving efforts to local forces and capacities



For HSI, MPICE data were then integrated into an analytical tool suite in which Logos analyzed the data to provide three unique outputs:

- ❖ comparative trend analyses between conflict drivers and institutional performance (is one rising or falling, relative to the other over time?)
- ❖ comparative trend analyses across the methodologies (are they indicating comparable outcomes over time?)
- ❖ comparative trend analyses of progress according to sector (is one sector progressing over another over time?).

In general, MPICE can be applied as an M&E system at the national level to regions that include parts of multiple countries (such as the Mano River Union in West Africa), and to focused, tactical areas of interest (such as Cité Soleil). The tailoring aspect of the MPICE development effort allows it to function in a full spectrum of scales. A Web-enabled tailoring wizard also allows users in different physical locations to narrow down over 600 built-in measures to suit their needs for a particular environment.

Fundamental to HSI's work is the ability to monitor efforts and evaluate progress toward desired outcomes, or goals, in Cité Soleil. To support HSI's efforts, it is necessary to collect baseline

economic, geographical, and sociological data for Cité Soleil and to track how this data changes over time.

One of the strengths of its M&E process was that at the outset, Logos worked with HSI to develop a strategy for data collection that best reflected desired program goals. This strategy was then used to frame the analysis plan. Following this, the M&E team developed an analysis plan that incorporated knowledge of the environmental context, availability of data, and the relative applicability of the existing and prototype data collection methodologies, such as the expert elicitation method or expert knowledge method. This plan was developed prior to each data collection, with intense local participation, and evolved based on lessons learned from prior work.

Sorting out what we learned by doing M&E in this environment using the MPICE system, we can put the lessons into two categories. The first is the program implementer's strategic and operational perspectives, as well as recommendations and lessons about doing evaluation. Second is the analyst's explanation and lessons on using the MPICE system.

Strategic Perspectives

Some recent research and much anecdotal evidence show that successful counterinsurgency efforts are really a combination of multiple efforts on a broad socioeconomic and military front. There is no magic bullet; many different things have to work right in the field for counterinsurgency efforts to reach the tipping point. The same is true for stabilization. An advantage of the MPICE approach is that it is not tied to any one program, or any one agency, and it is broad.

What follows are a number of lessons learned from applying MPICE. Overall, the

MPICE tool was both flexible and provided well-founded results, which appeared less impressionistic than most other systems used in Haiti.

Stabilization Is Not Development. Just because a program may be using traditional development tools and approaches, people fall into the mental trap of assuming that we need these traditional M&E tools to track development indicators. Using the MPICE framework's distinction between drivers of conflict and institutional performance was one way to clarify between doing a project to make people healthier or more educated, and doing a project to make a place calmer, safer, and more governable. The stabilization program might build schools or health clinics, thus achieving the outcomes of a development program, but for a stabilization program, those are tools (outputs) only. Those outputs may not be what we want to measure. We may want to measure the change in people's attitudes to local government, or the development of local leadership in carrying out the projects. Building a school is not good in and of itself, at least for a stabilization program. Stabilization outcomes may well be harder to measure, and harder to achieve, and a planner needs to be clear about the differences.

Discipline Is Good. Part of the value of the MPICE framework is deciding which specific goals and which tied measures are to be taken from the menu of more than 600. While seemingly mindless and rigid, this discipline is a plus for the implementer because it avoids "cherry picking" (that is, measuring only what might make the project look good). Any standard evaluation setup process requires that the operator follow a logical and defensible process. With this M&E program, it meant someone else had done the hard work, and we were not reinventing and then defending the wheel.

Be Careful What You Wish For. A corollary is that we do need to find goals and indicators that are actually measurable in the terrain. At the start of HSI, we passed around the goals and indicators list to multiple agency representatives and within the team itself. We selected measures, agreed across the agencies on the final set, and then set out trying to use them. Some of them did not work—such as the measure that asked how many people had bank accounts; the number was so low it was not measurable, and it did not indicate anything about the economic state of the individual or his trust in institutions, as there were no banks in the zone. If we were to go through this process again, we would use more local advisors earlier on to help decide the important measures, and even what goals to pick.

We did not use solely local input to decide what was important to measure because it is possible to imagine situations where the local interests are different from U.S. or international interest. Witness the case of working health clinics, which the locals might see as success, and the U.S. Government could see as irrelevant to stabilization.

Evaluation Can Be a “Forcing Function.” Getting agencies to work together is difficult, even at the best of times, and in a crisis environment, with little data and a lot of conflicting opinions, it is very difficult. Prior agreement on goals and indicators can help in getting people to aim at the same target. While the interagency process produces many cleared statements of goals, they are usually not actionable (that is, they cannot be broken down into clear plans). Democracy, stabilization, economic development all mean little when on the ground. Using a strong evaluation process to force agreement on specific indicators and goals can be valuable.

The 4 to 6 Percent Solution. With an always limited budget, a planner must decide

how much can be spent on evaluation. As a rule of thumb, a planner should assume that 4 to 6 percent of the budget should go to evaluation and collecting metrics. This is not much, but if it is not fenced off, it will quickly be raided as the budget is developed. As an incentive, keep in mind that if good metrics are proven and indicate program success, it is easier to get more funds in the future. If there are merely a few anecdotes or spotty and unreliable data, another program with solid metrics is more likely to win the next grant. The depth of data paid off for the HSI program. The Cité Soleil data later contributed to refunding the program to work in another part of the city.

Operational Perspectives

In moving from strategic plans to field operations, there is always the sound of grinding gears as an implementer tries to fit plans to reality. But even in the cases where MPICE did not work as well as we thought it would, it still worked better than the alternatives.

The U.S. Government Needs an Off-the-shelf Evaluation Capability. We have previously mentioned the value of having a common U.S. interagency and even internationally accepted capability for M&E of stabilization. That said, it was still difficult to launch this evaluation of our program and continue it. Contracting and deployment of evaluation were so long and drawn out that they affected results. In a crisis intervention, there is a need for a baseline study to be done simultaneously with the deployment of the stabilization team, or even before deployment if possible. Yet the United States has no accepted standard for what is needed in evaluation (as one example, the MPICE framework is still a draft) and has no way to contract for this service quickly.

View of Cité Soleil, Port-au-Prince's most notorious slum and focus of the Haiti Stabilization Initiative

State Department (Laurence Kent Jones)



HSI had enormous difficulty contracting for the original test. The baseline data were actually 6 months into the deployment (thus missing the improvements from the start of the program). For the expansion of HSI into the Martissant gang zone in January 2010, there is still no signed contract after many months of efforts, literally hundreds of e-mails, and the involvement of contracting authorities from two agencies and three different bureaus. In other words, rather than getting easier, things actually worsened the second time around, as different agencies became involved. This situation cries out for an Indefinite Delivery Indefinite Quantity contract done in advance and quickly available when necessary. The contract should provide a range of approved methodological tools and analysis techniques and provide skilled implementers for use in whatever part of the world is necessary.

M&E and MPICE Program Management Tools. The MPICE focus on measuring strategic or operational outcomes meant that it was primarily a strategic- or operational-level tool. In Haiti (apart from overall success measures), we also experimented with mining the data for program decisions and to measure project implementation results. It does provide detailed results at the different time slices because it uses different methodologies and allows different views of the same issues using different tools. However, it is not able to give much real-time feedback, and we could not easily separate outcome results to measure stabilization results of spending in education versus health, for instance. The major time lags made it difficult to react to new pressures or incidents using data from the M&E results. We recommend a quick and cheap spot survey mechanism for those burning questions that come up between phases.

Causality and Bang for the Buck. Another difficulty relates to the causality or firm attribution argument. Most stabilization efforts (and many development or security efforts) suffer from a



simple problem: proving that spending effort or money *here* equals a change in attitude *there*. As happens in social science, there is sometimes a demand that there be dependent and independent variables, which is sometimes an artificial way of examining intangible issues. While the MPICE framework makes it much clearer which is being measured (conflict drivers or institutions), there is still a final leap to showing that those indicators are being affected by the program. This is still better than many programs that make vague assumptions on theories of change and then cannot break down the process.

Ideally, we would have liked to have something akin to a control group and then run the same measures on one area that we were running in the project area. There were also ethical issues about repeatedly surveying a zone but not working in it. In reality, it was not easy for us to find a true equivalent area. In Haiti, we did not have a formal control group with which we could compare findings, but because the zone was essentially abandoned due to the security situation, we were fairly confident that any sectoral outcomes and impacts may have been due to the HSI effort, if only because there were no other significant actors in the zone. There were few other explanations for changes in measurements beyond the changes caused by our interventions. This would probably not be true in a national scale test, where there might be multiple international actors, not to mention nongovernmental organizations.

Exogenous Factors. For the implementer at either a local or national level, how can results be separated from background noise and how can what is happening on a national and international level be separated from what is happening in the zone? Exogenous factors played a big role in the measures in Haiti in the end. During the period when we were measuring, things visibly improved in

Cité Soleil and residents recognized the change. Unfortunately, at the same time, Haiti experienced a sharp rise in fuel prices, causing a series of national food riots in 2008 that toppled the government and left it in disarray for months. In 2009, Haiti had four tropical storms and hurricanes in 1 month. And in 2010, it suffered an earthquake that killed 300,000 people and left 1.3 million people homeless within minutes.

Apathy Toward the Overall Evaluation.

Within a multiagency team, despite the whole-of-government mantra, we found that some agencies have more interest in some things than in others. The overall program results may not be their main objective, especially if their agency measures the success of the program using a different yardstick. While not surprising, it does mean that their interest in the overall data is minimal. They just want their data to be good; the rest of it is meaningless to them. A good evaluation tool can serve a valuable “forcing function” for getting agencies to play together, but only if the agencies and actors believe they will be measured on the *overall* success, and not just, for instance, on how quickly they moved the money, or what was built, or how few crimes were reported. Combined with general resistance to being “evaluated to death,” this can be quite contentious when their part of the program is deprioritized for another part of the program.

If It Cannot Be Explained, It Never Happened. Any M&E program would benefit from the inclusion of more sophisticated analytical and visualization techniques (factor analysis, for example; video files, and so forth). In addition, as it currently stands, our analysis is good at providing cross-sector/driver to institutional analysis, but it is still relatively immature on how best to visualize or illustrate this analysis graphically. Indeed,

a picture tells a thousand stories. Our recommendation is to enhance the graphical capabilities in the MPICE tool and to increase visualization options beyond the standard bar chart and line plot functionality that currently exists. The need for pictures was made abundantly clear in the briefings in Haiti as people’s eyes glazed over while we explained the many valuable and redundant features of MPICE. There is a need to better visualize the complex analysis we provided to the U.S. Government as well.

Analytical Issues

Be Sure Everyone Is on the Same Page.

One of the most important challenges associated with the application of the MPICE framework—or even the variant of the framework that we applied—is the wide and varied understanding and definition of terms such as peace, conflict, and stabilization. Because MPICE is a multisector measures framework, its application should naturally be in a multiorganization environment, where experts and organizations with a stake in each sector of interest would participate. While the MPICE framework could be used by a single organization to do a multisector analysis of progress, the results of this application would likely be less rich and relevant than a multiorganization assessment, as no one organization can have a full appreciation of all of the sectors represented in MPICE. The challenges associated with multiorganization assessments are many, but for the purposes of MPICE, one of the most fundamental challenges will be that of agreeing on the meaning and importance of a term such as stabilization among the organizations involved in the assessment.

The MPICE framework is based on a theory that a reduction in the drivers of conflict,

combined with an equitable increase in the capacity of local institutions, eventually leads to stability. This theory, while logical, may not be applicable to all conflict situations. Furthermore, in a multiorganization assessment, there will likely be disagreement between participating organizations as to what theory or theories link to positive change in the area of interest. Development experts will likely lean toward theories based on long-term, sustainable development. Military planners will normally advocate the creation of a safe and secure environment and winning the hearts and minds of the people. Conflict experts may look to a number of different, interrelated theories regarding how to resolve conflict and bring about stability. This discussion of theories of change can have a significant impact on the application of the MPICE framework, especially given the organization of the framework around one particular theory.

How Much Is Enough? An additional challenge associated with applying the theory that a reduction in the drivers of conflict, combined with an increase in the capacity of local institutions, will eventually lead to stability is in setting thresholds of progress and/or success in relation to stabilization. How do we determine when the conflict drivers have been reduced enough to signal stabilization? How do we determine when institutional capacity has increased enough to signal stabilization? The simple diagram that is often shown during briefings about the MPICE framework depicts this theory with two arrows crossing each other on a single X–Y graph (see figure). This implies that these two different factors can be measured with the same units of measure and on the same scale, and that there are thresholds that indicate appropriate progress has been made.

Looking back over the data, we were more successful at pushing down the drivers of conflict than we were at pushing up the strength of the institutions. Again, this fits with what we observed over time. It was not that the gangs were strong; it was that the state was weak. However, at the time in the field, this was not so clear. We were making progress, but we did not really know how much further we had to go; there was no clear endstate. We did not know where the “X” was.

looking back over the data, we were more successful at pushing down the drivers of conflict than we were at pushing up the strength of the institutions

Decide What the Endstate Will Look Like and Be Sure That Everyone Else Has the Same Endstate. Several of the prototype applications of the MPICE framework that were conducted prior to and/or in tandem with the HSI M&E program arbitrarily identified progress toward stabilization without consideration of the U.S. Government mission’s own goals, processes, mandates, decisionmaking priorities, and problem-solving approaches. If any metric tool is to be formally adopted as a U.S.-wide M&E tool, this integration must take place as it was with the HSI M&E program.

A Good Datum Is Built, Not Found. The data analyzed were derived from several qualitative and quantitative data collection methodologies. By analyzing different data from different collection methodologies, we were able to build on the strength of each type of data collection and minimize the weaknesses of any

Boy watches as U.S. Sailors arrive at New Hope Mission in Bonel, Haiti, January 2010



U.S. Navy (Kristopher Wilson)

single methodology. This multimethod approach can increase both the validity and reliability of data. Quantitative and qualitative techniques provide a tradeoff between breadth and depth and between generalization capability and targeting to specific (sometimes limited) populations.

It is hardly surprising that in stabilization situations, quantitative data varies between weak and nonexistent. It was one part of the HSI three-legged data stool, but it was hard to find; a weak, underdeveloped government does not gather much data, and even less in dangerous zones. Other options to explore would be the use of microeconomic activity indicators gleaned from photographs or quick surveys.

In the case of Cité Soleil, much of the quantitative data that one might want to use is not available with the type of granularity needed when looking at just one small zone of the country. National crime statistics, for example, may not break down easily to a specific department, or school data may divide by education district, not by zone. This presented some challenges that were eventually overcome as we modified the framework and introduced greater flexibility. Similarly, across all five MPICE framework sectors, data availability, data reliability, and data accuracy problems arose and required significant adaptation of the framework and the Logos processes for data collection and analysis. These challenges led to the fact that only some of the data collection methodologies described in the framework provided good trend data over the three phases.

Ask the Community for Help in Evaluating Results. A community-based participatory approach is critical for a successful implementation of an M&E project. Ethnographic fieldwork among selected local communities combined with focus groups provided community perspectives and concerns

related to progress in Cité Soleil and the post-earthquake recovery process. Interviews were used to evaluate complex subject matter and to gather additional information in more detail from expert or high-status respondents or to discuss sensitive subject matter (such as criminal influences, corruption specifics) that is deemed inappropriate for survey/polling or even focus groups. We used interviews throughout the three phases to uncover inconsistencies between other data sources and to explore particular findings gathered from other methodologies.

How Quickly They Forget. Finally, as we went through iterations of the measures in the phases, we discovered that we were not getting the same data, or even variants of the same data. We thought we were winning, but the data did not show it. Why? In the final phase, we added a paired comparison of goals desired by both HSI and the community itself. This showed that in the first phase, security was seen as a crucial problem, and efforts to attack it were viewed positively. When security improved, it suddenly dropped off people's personal screens, and employment and education suddenly became the issues of concern. This priority change was reflective of the shifting goals of the community and the fact that these personal goals were not aligned to HSI's overall goal set.

In an ideal situation, we would have developed a separate set of questions (presented to community leaders in a controlled environment) that would have determined which goals were more critical relative to others by asking respondents to identify and rank which issues were important. This would have helped us to assign weightings that were more accurate to each measure and goal for our mission. For example, an HSI survey can ask respondents how satisfied they are with their electricity

access and the condition of their roads. In this case, let us assume that 90 percent are satisfied with the roads and that 10 percent are satisfied with their electricity access. These results will not affect how we weight the importance of each measure, but a separate questioning process or structured "pairwise" process asking key leaders or other people which was more important to their quality of life, roads or electricity, might affect how we weight our goals, and therefore the inputs we use to achieve those goals.

Assigning values or weights to measures, indicators, or goals is also a critical step in the analysis process. It allows the policymaker, decisionmaker, or analyst to designate the relative importance of one finding against another. Depending on the issues driving the conflict and the role that institutions have played in exacerbating rather than resolving conflict, some indicators may be more salient. Weightings on a scale of 0 to 1 may have security-related measures, indicators, or goals weighted heavier than economic or social well-being measures, indicators, or goals. The analyst cannot assign value; it must be assigned through a consultative process with subject matter experts, decisionmakers, and policymakers. We proposed, but did not fully execute, a technique designed to weight the M&E data responses.

Conclusions

In the same GAO review of stabilization evaluation and monitoring, the report ends with a recommendation: "We have previously reported that key practices for enhancing and sustaining interagency collaboration include developing mechanisms to monitor, evaluate, and report the results of collaborative programs."²

Applying the MPICE framework along with the multiple data collection methodologies and analytical techniques was fruitful and

provided the program implementers with the best data possible in the difficult and deteriorating environment of Haiti. Measuring progress in a conflict environment is always a challenge, and even with a serious effort using sophisticated M&E methods, analytical techniques, and tools, including the MPICE framework, our program produced almost as many questions as it answered. We improved our efforts over each phase, and presumably, if we had had more than three collection phases (or maybe just one less earthquake), we would have had far more data to analyze and use for planning.

Regardless of the results, for planners thinking of future applications, the importance of planning for evaluation from the beginning and designing the stabilization program with that in mind is crucial, and there is a clear need for continued improvements in the tools and their visualization. Even clearer is the need to improve contracting for evaluation programs so that proper baselines and ongoing data are collected. Most importantly, a good monitoring and evaluation plan, in highlighting the theory of change in core assumptions in the stabilization program, can serve to concentrate the focus of many different organizations, clarify the strategy, set objectives, and guide tactics. This is valuable even before the evaluation results are in. **PRISM**

Notes

¹ U.S. Government Accountability Office (GAO), *International Security: DOD and State Need to Improve Sustainment Planning and Monitoring and Evaluation for Section 1206 and 1207 Assistance Programs*, Report to Congressional Committees, GAO 10-431 (Washington, DC: GAO, April 15, 2010).

² Ibid.