Monitoring and Evaluation of Department of Defense Humanitarian Assistance Programs

Colonel Eugene V. Bonventre, U.S. Air Force

If you took all the bricks in the Pentagon and laid them end to end, they would reach around the earth four and a half times.

—Pentagon tour guide, 7 December 2006

Colonel Gene Bonventre, U.S. Air Force, is senior international health policy advisor in the Office of the Deputy Assistant Secretary of Defense for Partnership Strategy. He served with Joint Forces Command from 2004-2006 and with humanitarian assistance teams in Sierra Leone and Pakistan. COL Bonventre planned and helped execute 30 humanitarian assistance exercises in 16 countries and has deployed to 6 countries. He earned a B.S. from the University of Notre Dame and an M.D. from Loyola-Stritch School of Medicine. A board-certified general surgeon and flight surgeon, COL Bonventre is also a graduate of the Air Command and Staff College and the Air War College.

PHOTO: Two young Afghan boys overlook a USAID/Afghanistan-funded medical clinic in Wardak Province, 2005. The project was implemented by USAID partner "Shelter for Life." (Shelter for Life International, William R. Billinglsey)

THE U.S. DEPARTMENT OF DEFENSE (DOD) executes humanitarian activities primarily through the Overseas Humanitarian, Disaster and Civic Aid (OHDACA) program. The OHDACA program includes three sub-activities: the Humanitarian Assistance (HA) program, the Humanitarian Mine Action program, and Foreign Disaster Relief and Emergency Response. Activities funded by the OHDACA appropriation are intended to mitigate the effects of natural and man-made disasters, to shape the environment in which DOD operates by providing access to critical areas and by influencing civilian populations, and to improve the capacity of vulnerable nations to better prepare for disasters. The ultimate beneficiary of OHDACA activities is the civilian population, and the activities should always have an appropriate and positive influence. For instance, renovating a school should positively impact primary education, and renovating a clinic should positively impact the civilian health sector.

Civilian U.S. Government agencies evaluate the effectiveness of their programs through monitoring and evaluation (M&E), but equivalent analyses of DOD humanitarian assistance programs have been either ad hoc or entirely lacking.[1] "Monitoring" is the ongoing, systematic collection, analysis, and use of data during the course of a project.[2] "Evaluation" is the periodic review of program activity, outcome, and impact, with an emphasis on lessons learned.[3] This article presents the case that DOD should institute both monitoring and evaluation of HA activities in order to assess their effectiveness.

The "How" and "Why" of Measuring HA

Every organization currently involved in humanitarian assistance faces the challenge of how to measure the impact of its work. Despite nearly 40 years of experience in M&E, the United States Agency for International Development (USAID) still struggles to quantify, and demonstrate to decision-makers, the impact that its programs have. While DOD has extensive experience with battle damage assessment, its M&E methods for humanitarian assistance are in their infancy. The Pentagon has instituted "measures of effectiveness" (MOE) for virtually every DOD program but HA. The Defense Department need not develop monitoring and evaluation methods in a vacuum, however. USAID's several decades of experience is a great start point. Other agencies' experiences and lessons learned can likewise serve as a base for development of M&E techniques.

66   January-February 2008 ● MILITARY REVIEW

Why should DOD measure the impact of HA programs? There are several important reasons. First, doing so can allow planners to make mid-course corrections on current projects, and it can provide them with information to improve the quality of future activities. By creating a feedback loop of lessons learned, the M&E process in HA would improve efficiency and ensure that projects contribute to operational objectives. Planners could then emphasize activities that are more cost-effective, which is especially important because every year the number of projects that combatant commands apply for exceeds the funds available. Second, collecting and sharing data would increase planners' ability to deconflict activities with other agencies and non-governmental organizations (NGOs). Third, data analysis helps to showcase quantifiable results, thereby minimizing the chances of negative press surrounding HA activities.[4]

But most importantly, DOD should measure HA programs because transparency is a core strength of our democracy. Groups like Hezbollah, Hamas, and even Al-Qaeda engage in prima facie humanitarian and social service activities, but ultimately their true motivations become apparent: the manipulation of people toward violent ends. In contrast, DOD's humanitarian programs should have a demonstrably quantifiable humanitarian impact.[5] Since terrorist organizations will usually be able to act more quickly than DOD (because they are not impeded by bureaucracy, ethical norms, and legal restrictions), any demonstrable positive benefit to the civilian sector offers DOD the chance to prevail in the long term over extremist propaganda.

As a point of contrast, in Vietnam, DOD spent $500-$750 million on MEDCAPS (medical civic aid programs) that provided medical care to 40 million civilians. However, in the absence of data, analysts have failed to reach any significant conclusions about the results of those programs.[6] Now is the time to avoid having to encounter the same situation in the future, since we stand to gain much by accurately assessing our HA activities.

> **DOD should measure HA programs because transparency is a core strength of our democracy.**

DOD's initial attempts at measuring HA effectiveness will very likely be less than perfect. But merely attempting to quantify results will gain the department credibility. The resulting goodwill and improved civil-military relationships may even result in cooperation to refine future evaluation efforts. Regardless, the only way to ensure that we have long-term access to the areas we can affect, far beyond the short period of time that DOD personnel are on the ground, is to ensure that good civil-military relations continue. If the population feels abandoned at the completion of a project, all will have been for naught. For similar reasons, the best way to ensure positive influence is to quantify the benefits that the civilian population enjoys as a result of a given project and then feed that information back to the host nation.

Common Terminology

Because agencies often use the same words to mean different things, any discussion of monitoring and evaluation requires a common understanding of terms. The definitions in this article were taken from sources within DOD and other agencies:

● A "standard" is a reference point that allows comparisons. It is a set of criteria, guidelines, or best practices. The SPHERE Project publishes a handbook of minimum standards in humanitarian assistance and disaster response commonly used by civilian agencies.[7]

● A "goal" is an overall statement of intent. It is broad, timeless, and unconcerned with particular achievement within a specified time period.

● An "objective" is exactly what will occur, how it will be accomplished, and to what standard of performance.

● "Indicators" are quantitative or qualitative measures of standards and are used to correlate or predict the value or measure of a mission, program, system, or organization.[8] Indicators should have "SMART" (specific, measureable, achievable, relevant, and time-bound) characteristics.

"Specific" means focusing on a narrowly defined aspect of a unit's mission. "Measurable" means showing progress and providing data for mid-course adjustments and improvements. "Achievable" relates to focusing on realistic targets rather than vague end-states. "Relevant" indicates that a strategic goal, major initiative, or core service has

been measured. "Time-bound" means it applies to a specific time frame. All indicators should complement one another. The SPHERE handbook contains a variety of indicators that might be helpful to military planners.

● "Performance indicators" (also known as "process indicators" or "achievement indicators") describe the output of an activity or how well that activity functioned. Performance indicators are important to measure, but they don't tell the whole story.

● "Outcome indicators" (also called "impact indicators") measure the extent to which an activity contributed to the overall goals of a program.

● "Measures of effectiveness" are combinations of key indicators from multiple sectors (or functions) used to determine overall progress toward attaining mission objectives. In joint doctrine, MOE are well described for large operations, and they are often used to determine transition strategies or redeployment milestones.[9] In the case of HA activities, MOE should measure access, influence, sectoral impact, and capacity building.

● "Baseline data" are measures of specific indicators that exist prior to project implementation. The partner nation, USAID, international organizations, or NGOs will usually have baseline data, although it may be very limited in conflict zones. If no baseline data exists, it may be prudent to collect it at the beginning of large projects.

To illustrate how these terms should be used, a project could be designed with a *goal* to improve village health. A possible *objective* would be to distribute, within one deployment, mosquito bed-nets to 95 percent of the village's residents. A *process indicator* would be the number of bed-nets distributed. An *outcome indicator* would be the percentage of villagers who actually use bed-nets. And a *measure of effectiveness* would be the decrease in the number of new cases of malaria in that village. Some indicators can be measured during or immediately after the project, and others will require one or more follow-up visits.

Roles and Perspectives

Project designers should monitor individual projects by developing both performance indicators and outcome indicators. In many instances, designers can tap into existing sources for baseline data and specific indicators. The host nation is usually the

> *Project designers should monitor individual projects by developing both performance indicators and outcome indicators.*

ultimate source, since most governments, even in resource-poor countries in conflict, will have some idea of what data they should collect and what has been collected in the past. Moreover, the host nation should be involved in, central to, and ultimately the owners of, every DOD HA project.

However, an easier source to access for data collection is the in-country or regional USAID office. USAID is not always an integral member of the country team at the American embassy, so planners may have to seek them out. A pre-planning conversation with USAID can often be a one-stop shopping event in which project designers access international organizations and NGOs in the sector of interest and identify competent and talented host-nation personnel. When using these sources, it is far better to collect your own data in the same format as the host nation or USAID than to create an ad hoc system.

In the data collection process, project officers measuring an impact on specific HA sectors or functions should also attempt to measure indicators that relate to DOD-specific goals. For instance, an HA project might be designed to improve a nation's capability to respond to a disaster or to an outbreak of pandemic influenza, and that could factor into Soldier readiness. Similarly, if a school is to be renovated, project goals may include "positively influencing a village that is prone to insurgent manipulation." Project designers in such circumstances may need to create new measures to look at these DOD-specific goals, but even then, examples can be gleaned from sources outside DOD.

Although DOD-specific goals would be important to the combatant commander, he need not be involved in data collection, except perhaps to share lessons with other project designers.

Combatant commanders' HA managers should provide oversight of individual projects, but should also be interested in evaluating overall programs. A "program" might be a multiple-year series of

related projects, each of which individually might not directly achieve theater goals, but taken together should contribute to realization of the theater strategy. Such a series of related projects might warrant development of specially tailored MOE. This process is usually labor-intensive, but it can be made easier if individual project metrics are designed to feed into them. A strategically holistic approach from the beginning would help designers implement that process. The combatant commander could transmit the resulting qualitative information to DOD senior leadership, describing program impacts that are difficult to quantify. Security cooperation assessments serve this same purpose for military-to-military activities.

Data collected should actually affect decision-making. The data reported by the tour guide at the beginning of this article is interesting, but what decisions would be changed by knowing the number of bricks in the Pentagon? Before collecting data, project planners should consider the target audience whose decisions rest on what is collected. For HA activities, the primary user of the data is the project designer, who needs to know if the activities he or she planned actually had the intended effect, so that they can improve the planning process for future activities. The designer can then summarize the data and pass it along to higher headquarters, where it can be used by program managers to assess the effectiveness of the whole program.

Partnerships

DOD cannot, and should not, monitor and evaluate humanitarian assistance missions in isolation. Certain goals of HA activities overlap with goals of other agencies. Both the Office of Foreign Disaster Assistance and DOD, for instance, have an interest in improving host-nation capacity to respond to disasters. USAID and DOD may each want to make a positive impact in the health sector in an area vulnerable to extremist influence. Since DOD HA activities must be done in partnership with the host nation, some of DOD's and the host's goals presumably should overlap. Therefore, rather than duplicate effort, a project designer's first step should be to find out if the host nation already collects similar data. If it does, the host's data may serve as a guide for formatting other or additional data collection, and it may provide baseline data for comparison.

USAID, Ben Barber

Workmen at the Panjsad Family High School in Kabul, Afghanistan, battle the cold but keep on with renovations under a USAID grant, preparing for the return of 10,000 students after winter break, 21 January 2006.

The designer's second step should be to query other agencies, which normally require their implementing partners (NGOs and contractors) to collect data. For health projects, the Uniformed Services University and the Center for Disaster and Humanitarian Assistance Medicine have operational expertise and can provide advice and support to both project designers in the field and combatant commanders' HA managers. Occasionally, one can find common ground for disaster preparedness with organizations outside the U.S. Government. For instance, the UN International Strategy for Disaster Reduction may already measure a country's disaster response capabilities.

Yet another data source is academia. Johns Hopkins University developed a "balanced scorecard" method for assessing the capabilities of health

DOD cannot, and should not, monitor and evaluate humanitarian assistance missions in isolation.

facilities in Afghanistan.[10] Harvard, Tulane, and the University of South Florida have significant experience in program and project design and evaluation. Leveraging any of these efforts will save DOD time and money, and it will facilitate data-sharing with other agencies and organizations. This sharing also increases DOD's credibility with other agencies and allays fears that the department is encroaching on other agencies' territory.

Where DOD goals are unique, project designers will need to develop their own indicators. Development from whole cloth will likely be the case at the program level if DOD chooses to quantify the impact of HA programs on "access" and "influence." However, even this process can benefit from methodologies already well-developed in other agencies and academia. Some quantitative analogy somewhere will almost certainly serve as a heuristic for developing new indicators.

Project monitoring is best done in-country through a collaborative effort with the country team and USAID. Program evaluation is best done at the combatant command level—preferably periodically, and possibly by an independent contractor. Given the significant number of HA projects implemented by contractors, precautions ought to be taken to avoid having contractors evaluate each others' efforts, especially when they are competing or belong to the same company. Academic institutions or NGOs could have a role at either level, although they may bring their institutional biases to evaluations. Nonetheless, they can be a cost-effective way of adding an independent, and possibly more credible, view from outside.

Resources

According to both the Department of State (DOS) and USAID, collection and analysis of data usually consumes from 8 to 10 percent of a project's total cost. These numbers provide a reasonable starting point for planning purposes. In developing an overall budget for worthwhile M&E in HA projects, combatant command HA managers may choose to develop a pilot project. Measuring a small number of moderate-sized HA projects in secure, accessible areas may be the best option. Projects undertaken in well developed civil societies with potential academic partners, some willing personalities in USAID, and relatively little corruption in the host-

nation government would seem to offer the best possibilities. Pilot projects in developed countries like South Africa or other stable developing countries might demonstrate the utility of M&E for humanitarian assistance. The present OHDACA appropriation should be adequate to start this process, and if the process succeeds, it can be used to demonstrate the need for additional funds from Congress. The ultimate goal is a wider implementation of project M&E and a subsequent increase in the number of projects funded.

Occasionally, combatant commands fund HA activities in a "tier 3" (lowest priority) country simply because there is no other significant U.S. activity there, or just because they want to spread their influence throughout their area of responsibility. A recent review of medical projects conducted during a 12-month period in all combatant commands found that fully two-thirds of project proposals were in tier 3 countries. This only makes sense if resources are virtually unlimited and tier 1 (highest priority) and tier 2 countries are completely saturated with activities. Neither is likely to ever be the case for OHDACA. There may be good reasons for doing occasional activities in tier 3 countries, but they should be the exception, not the rule. A Defense Security Cooperation Agency review of HA activities in tier 3 countries would indicate the magnitude of the current involvement, and it may identify a potential source of funds that would be better spent on monitoring and evaluation.

The Example of Afghanistan

The following project, though running into challenges in the execution phase, nonetheless shows that it is relatively painless to develop interagency indicators that quantify the effects of DOD humanitarian activities on stability and security.

In 2006, Afghanistan's Minister of Public Health (MoPH) noted that the people of the Nuristan, Kunar, and Laghman border provinces routinely cross over into Pakistan, ostensibly to receive healthcare unavailable in their own provinces. During these excursions, they take drugs into Pakistan and return with guns. The provinces were too unstable to permit NGOs to establish enough clinics, so the minister asked the Combined Security Transition Command (CSTC-A) leadership for assistance in fielding mobile health clinics staffed

by local Afghan personnel. CSTC-A funded an Afghan NGO, Sozo International, which got buy-in and security guarantees from local tribal leaders in exchange for hiring local personnel. Sozo agreed to provide the provinces with medical training and free medical care. CSTC-A coordinated this with the in-country USAID team and the European Commission (both being major donors to the Afghanistan health system). The Secretary of Defense's Partnership Strategy office coordinated with the State Department Afghanistan desk, USAID's Asia Near-East Bureau, Health and Human Services, and the Uniformed Services University of the Health Sciences (USUHS) to select a menu of health indicators based on the "Basic Package of Health Services for Afghanistan."[11] MoPH developed the package with USAID assistance, so the menu contains no new indicators.

Since MoPH requires all NGOs in Afghanistan to collect these metrics, Sozo International is comfortable with the process and with sharing data with USAID, MoPH, and the European Commission. MoPH will provide baseline data. Existing security indicators, already measured by CSTC-A with Joint Staff (J-5) assistance, will be compared in three provinces with clinics and three without. Project funds will be used to add several questions about

…even in a complex operation in an insecure environment, one can design M&E without undue burden.

the clinics to a monthly population attitude survey conducted by a Kabul-based organization. Attitudes in the three provinces with mobile health clinics will be compared to attitudes in three adjacent provinces without the clinics. The implementing NGO hired a project manager to oversee metrics. CSTC-A mentors this person, collects security indicators, and provides progress briefs to senior leadership. Data will be used to improve the quality of healthcare in less secure provinces, garner the cooperation of other stakeholders, determine whether similar projects should be launched in other provinces, and determine whether or not additional funding is justified.

The mobile health clinics were scheduled to begin operation in September 2007, but have been delayed due to changes in the tactical situation; M&E results will be published when available. Since this is a large project ($1.25 million to field three clinics) in a sensitive area within a named operation, a formal, government-wide method of metrics development, including partnering with academia, was justified. Nonetheless, the M&E portion was designed by a USUHS student (a U.S. Air Force captain) with mentorship from the Partnership Strategy office (the author). Support from every relevant U.S. government agency was obtained with a few phone calls and two meetings. Since the project used health indicators previously developed by MoPH and USAID, local buy-in was straightforward. Tapping into existing population surveys and security indicators will quantify the impact that health activities have on security. This example demonstrates that even in a complex operation in an insecure environment, one can design M&E without undue burden.

A U.S. Army doctor speaks to a patient through an interpreter at a special clinic set up by the provincial reconstruction team in Ghazni Province, Afghanistan, 2006.

USAID, Ben Barber

Training

Although the principles outlined here are not complicated, they are unfamiliar to most military operators. Formal education and training on monitoring and evaluation would be time well-spent, and it would pay dividends down the road in a military officer's career. Virtually every DOD directive that mandates new tasks and responsibilities also requires use of measures of effectiveness. Perhaps "M&E" will ultimately become a mission-essential task for which DOD personnel will receive baseline formal education and routine refresher training. Until that time, however, conducting basic M&E for HA activities is bound to yield benefits.

Combatant command humanitarian assistance managers meet annually in Washington, D.C. for a conference sponsored by the Defense Security Cooperation Agency and funded by the OHDACA account. We hope to devote half a day to formal education on M&E techniques using basic principles taught by DOS and USAID, but modified to include DOD-specific requirements. DOS and DOD personnel or contractors could conduct this type of training, or one of several academic institutions could do it. Attendance should be required for all new DOD HA managers and be optional for current HA managers, some of whom could also serve as faculty. A similar course, funded by OHDACA, could be held at annual HA conferences in each combatant command.

For refresher training, M&E should be built into selected joint and service exercises that contain a humanitarian assistance scenario, particularly if other government agencies are involved. Such training would also help refine and customize theater M&E techniques, enhance familiarity with existing databases, and train a larger number of project officers.[12]

Conclusion

The complexity of today's security environment requires a new, sophisticated analysis of the efficacy of DOD humanitarian assistance programs. Assumptions should be replaced by formal attempts to quantify the effects of HA projects. That such measurements will never be perfect and causal relationships will never be definitively proven should not preclude attempts to develop practical assessment techniques. Merely attempting to quantify what has previously been thought unquantifiable will pay dividends in the quality of project design

USAID-supported health care clinics provide medical care to mothers and children, Afghanistan, 2005.

and implementation. Even if such attempts fail to achieve perfection, they will increase the credibility of DOD HA programs. Moreover, the interagency cooperation necessary for such a process will increase each agency's knowledge of the other agencies' principles and techniques and take them a step closer to a holistic, government-wide approach to addressing critical issues. ***MR***

NOTES

1. Jeff Drifmeyer and Craig Llewellyn, "Measuring the Effectiveness of DOD Humanitarian Assistance," Center for Disaster and Humanitarian Assistance Medicine, 1999; Doug Lougee, "Can We Build a Better Medical Civic Assistance Program? Making the Most of Medical Humanitarian Civic Assistance Funding," *The DISAM Journal* (February 2007); Derek Licina and Ken Schor, "Developing a Monitoring and Evaluating Capacity for the U.S. Department of Defense Humanitarian Assistance Program," *Military Medicine* 172, no.4 (April 2007): 339-45.

2. Office of Foreign Disaster Assistance, *Foreign Operations Guide,* version 4.0, September 2005.

3. Ibid.

4. CNN, "Audit: U.S. Lost Track of $9 billion in Iraq funds," 31 June 2005; ABC News, "$80 million Believed Wasted in Iraq Police Training Camp," 30 January 2007; Voice of America, "Officials: Billions in U.S. Iraq Reconstruction Money Wasted," 19 January 2007.

5. Even though DoD HA activities are carried out to gain access and influence, activities are also intended to benefit the civilian population. This distinction is not without controversy: by the internationally promulgated definition of "humanitarian," activities must take into account only the needs of the affected populations, without ulterior motives. For this reason, a more rigorous nomenclature for activities carried out as HA would be "civil-military operations," since additional motivations include attaining DOD operational and strategic objectives. However, the "benefit" to the civilian population must always be present as well.

6. Robert Wilensky, "The Medical Civic Action Program in Vietnam: Success or Failure?" *Military Medicine* 166, no.9 (September 2001): 815-19.

7. The SPHERE Project, Humanitarian Charter and Minimum Standards in Disaster Response (Geneva: The SPHERE Project, 2004), <www.sphereproject.org> (30 July 07).

8. Frederick (Skip) Burkle (Captain, U.S. Navy, Retired) has written extensively about measures of effectiveness and monitoring and evaluation. See Skip Burkle and P.G. Greenough, "Measures of Effectiveness in Disaster Management," in *Disaster Medicine,* 3d edition, ed. G.R. Ciottone (Philadelphia: Mosby/Elsevier Publishers, 2006), 333-35.

9. Joint Publication 3-07.6, *Joint Tactics, Techniques and Procedures for Foreign Humanitarian Assistance*, 15 August 2001, and Joint Publication 3-57, *Joint Doctrine for Civil-Military Operations*, 8 February 2001, page III-9 (both are in revision).

10. David Peters, Ayan Ahmed Noor, et al., "A balanced scorecard for health services in Afghanistan," *Bulletin of the World Health Organization* 85, no.2 (February 2007): 146-51.

11. "A Basic Package of Health Services for Afghanistan," Islamic Republic of Afghanistan, Ministry of Public Health, 2005.

12. One example is the Centre for Research on the Epidemiology of Disasters database (based in Brussels) used by the State Department. See <www.cred.be>.