

Free Executive Summary

Improving Democracy Assistance: Building Knowledge Through Evaluations and Research

Committee on Evaluation of USAID Democracy Assistance Programs, National Research Council

ISBN: 978-0-309-11736-4, 366 pages, 6 x 9, paperback (2008)



This free executive summary is provided by the National Academies as part of our mission to educate the world on issues of science, engineering, and health. If you are interested in reading the full book, please visit us online at <http://www.nap.edu/catalog/12164.html>. You may browse and search the full, authoritative version for free; you may also purchase a print or electronic version of the book. If you have questions or just want more information about the books published by the National Academies Press, please contact our customer service department toll-free at 888-624-8373.

This executive summary plus thousands more available at www.nap.edu.

Copyright © National Academy of Sciences. All rights reserved. Unless otherwise indicated, all materials in this PDF file are copyrighted by the National Academy of Sciences. Distribution or copying is strictly prohibited without permission of the National Academies Press <http://www.nap.edu/permissions/>. Permission is granted for this material to be posted on a secure password-protected Web site. The content may not be posted on a public Web site.

SUMMARY

Background

Over the past 25 years, the United States has made support for the spread of democracy to other nations an increasingly important element of its national security policy. Many other multilateral agencies, countries, and nongovernmental organizations (NGOs) also are involved in providing democracy assistance. These efforts have created a growing demand to find the most effective means to assist in building and strengthening democratic governance under varied conditions.

Within the U.S. government the U.S. Agency for International Development (USAID) has principal responsibility for providing democracy assistance. Since 1990, USAID has supported democracy and governance (DG) programs in approximately 120 countries and territories, spending an estimated total of \$8.47 billion (in constant 2000 U.S. dollars) between 1990 and 2005. The request for DG programs for fiscal year 2008 was \$1.45 billion, which includes some small programs in the U.S. Department of State.

Despite these substantial expenditures, our understanding of the actual impacts of USAID DG assistance on progress toward democracy remains limited—and is the subject of much current debate in the policy and scholarly communities. Admittedly, the realities of democracy programming are complicated, given the emphasis on timely responses in politically sensitive environments and flexibility in implementation to account for fluid political circumstances. These realities pose particular challenges for the evaluation of democracy assistance programs. Nonetheless, USAID seeks to find ways to attempt to determine which programs, in which countries, are having the greatest impact in supporting democratic institutions and behaviors and how those effects unfold. To do otherwise would risk making poor use of scarce funds and to remain uncertain about the effectiveness of an important national policy.

Yet USAID's current evaluation practices do not provide compelling evidence of the impacts of DG programs. While gathering valuable information for project tracking and management, these evaluations usually do not collect data that are critical to making the most accurate and credible determination of project impacts—such as obtaining baseline measures of targeted outcomes before a project is begun or tracking changes in appropriately selected (or assigned) comparison groups to serve as a control or reference group.

USAID has been seeking better evidence for the effects of its DG projects. In 2000 the Office of Democracy and Governance created the Strategic and Operational Research Agenda (SORA). Under SORA, USAID has commissioned studies of its DG evaluations and underwritten a recent cross-national study of the effects of its democracy assistance programs since 1990. A very encouraging finding from that study is that democracy assistance *does* matter for democratic progress. The study (Finkel et al 2007; see also the second-phase study, Finkel et al 2008) found that, when controlling for a wide variety of other factors, higher levels of democratic assistance are, on average, associated with movement to higher levels of democracy. These results provide the clearest evidence to date that democracy assistance contributes toward achieving its desired goals.

Unfortunately, it is also true that in a number of highly important cases—such as Egypt and post-Soviet Russia—large volumes of democracy assistance have yielded disappointing results. In addition to knowledge about general effects, USAID needs to know the positive or negative effects of specific projects and why DG assistance has been more successful in some contexts than in others. SORA turned to the National Research Council (NRC) for assistance in how to gain greater insight into which democracy assistance projects are having the greatest impacts. This report is intended to provide a road map to enable USAID and its partners to build, absorb, and act on improved knowledge about assisting the development of democracy in a variety of contexts.

Charge to the Committee

The USAID Office of Democracy and Governance asked the NRC for help in developing improved methods for learning about the effectiveness and impact of its work, both retrospectively and in the future. Specifically, the project is to provide:

1. A refined and clear overall research and analytic design that integrates the various research projects under SORA into a coherent whole in order to produce valid and useful findings and recommendations for democracy program improvements.
2. An operational definition of democracy and governance that disaggregates the concept into clearly defined and measurable components.
3. Recommended methodologies to carry out retrospective analysis. The recommendations will include a plan for cross-national case study research to determine program effectiveness and inform strategic planning. USAID will be able to use this plan as the basis of a scope of work to carry out comparative retrospective analysis, allowing the agency to learn from its 25 years of investment in DG programs.
4. Recommended methodologies to carry out program evaluations in the future. The recommendations for future analysis will focus on more rigorous approaches to evaluation than currently used to assess the impact of democracy assistance programming. They should be applicable across the range of DG programs and allow for comparative analysis.
5. An assessment of the feasibility of the final recommended methodologies within the current structure of USAID operations and defining policy, organizational, and operational changes in those operations that might improve the chances for successful implementation.

Overall Research and Analytic Design

In response to the first charge, the committee unanimously recommends a four-part strategy for gaining increased knowledge to support USAID's DG policy planning and programming. These are:

Recommendation 1: *Undertaking a pilot program of impact evaluations designed to demonstrate whether such evaluations can help USAID determine the effects of its DG projects on targeted policy-relevant outcomes.* A portion of these impact evaluations should use randomized designs since, where applicable and feasible, they are the designs most likely to lead to reliable and valid results in determining project effects and because their use in DG projects has been limited. USAID should begin the pilot program by focusing on a few widely used DG program categories. The pilot evaluations should not supplant current evaluations and assessments, but impact evaluations could gradually become a more important part of USAID's portfolio of monitoring and evaluation (M&E) activities as the agency gains experience with such evaluations and determines their value. (See Chapters 5 through 7 for a discussion of impact evaluations and how they might be applied to DG projects and Chapter 9 for the committee's recommendations.)

Recommendation 2: *Developing more transparent, objective, and widely accepted indicators of changes in democratic behavior and institutions at the sectoral level*—that is, at the level of such sectors as the rule of law, civil society, government accountability, effective local government, and quality of elections. Current aggregate national indicators of democracy, such as Freedom House or Polity scores, are neither at the right level for identifying the impacts of particular USAID DG projects nor accurate and consistent enough to track modest or short-term movements of countries toward or away from greater levels of democracy. (See Chapter 3.)

Recommendation 3: *Using more diverse and theoretically structured clusters of case studies of democratization and democracy assistance to develop hypotheses to guide democracy assistance planning in a diverse range of settings.* Whether USAID chooses to support such studies or gather them from ongoing academic research, it is important to look at how democracy assistance functions in a range of different initial conditions and trajectories of political change. Such case studies should seek to map out long-term trajectories of political change and to place democracy assistance in the context of national and international factors affecting those trajectories, rather than focus mainly on specific democracy assistance programs. (See Chapter 4.)

Recommendation 4: *Rebuilding USAID's institutional mechanisms for absorbing and disseminating the results of its work and evaluations, as well as its own research and the research of others, on processes of democratization and democracy assistance.* In recent years, USAID has lost much of its capacity to assess the impact and effectiveness of its programs. Without an active program of organizational learning so that senior personnel and DG officers have structured opportunities to discuss the results of pilot evaluations, compare their experiences with DG programs, and discuss the research carried out by USAID and especially other

scholars, implementers, and donors, the fruits of the committee's first three recommendations will not be usefully integrated with the experience of DG officers in a way that will improve DG program planning, design, and outcomes. (See Chapters 8 and 9.)

Discussion and Strategies for Implementation

The following sections provide more detail on the reasons behind these recommendations and discuss organizational issues at USAID that will affect the agency's ability to implement them.

Recommendation 1: Undertaking a Pilot Program of Impact Evaluations

Charges 4 and 5 asked the committee to recommend methodologies for future program evaluations and to evaluate their feasibility. These issues are addressed first, however, because the committee believes that, among the charges it was given, improving USAID's ability to more precisely ascertain the effects of future DG programs has more potential to build knowledge of what works best in DG programming than either retrospective analyses (given the limits found in the collection of data on past DG projects) or improving the definition of democracy. The committee thus investigated USAID's current evaluation methods and explored a range of designs for improved evaluations that could be applied to DG projects. The committee also commissioned teams of consultants to visit three diverse missions—in Albania, Peru, and Uganda—to assess the feasibility of applying those designs—in particular impact evaluations—to actual ongoing or planned DG projects. Of course, these evaluations, like all of USAID's evaluations and research, must be part of a broader learning strategy if the agency is to benefit; these organizational aspects are discussed separately below.

What Are Impact Evaluations?

Most current evaluations of USAID DG projects, while informative and serving varied purposes for project managers, lack the designs or data needed to provide compelling evidence of whether those projects had their intended effects. An *impact evaluation* aims to separate the effects of a specific DG project from the vast range of other factors affecting the progress of democracy in a given country and thus to make the most precise and credible determination of how much DG projects contribute to desired outcomes.

As the committee uses the term, what distinguishes an impact evaluation is the effort to determine what would have happened in the absence of the project by using comparison or control groups, or random assignment of assistance across groups or individuals, to provide a reference against which to assess the observed outcomes for groups or individuals who received assistance. Randomized designs offer the most accuracy and credibility in determining program impacts and therefore should be the first choice, where feasible, for impact evaluation designs. However, such designs are not always feasible or appropriate, and a number of other designs also provide useful information to determine the impact of many different kinds of assistance projects. For example, when there is only one group or institution receiving assistance, comparisons may be made across time by using a set of carefully timed measures before and after the project while controlling statistically for long-term trends or key events. Impact evaluations are designed according to standard protocols of evaluation research; yet the choice of a particular design and decisions about how to adapt the design to a particular project require skilled craftsmanship as much as science.

Current Approaches to Evaluation in USAID

The committee's review of current approaches to the evaluation of development assistance in general, and USAID DG programs in particular, found that:

- Very few of the evaluations undertaken by international or multilateral development and democracy donors are designed as impact evaluations. There are signs that this is changing as some donors and international agencies are beginning to implement new approaches to evaluation. The Millennium Challenge Corporation and the World Bank in particular have undertaken efforts to increase the use of randomized designs in evaluations of their economic assistance and anticorruption projects. A few NGOs also have undertaken randomized impact evaluations of their democracy assistance efforts.
- Within USAID the number of evaluations has declined for *all* types of assistance programs. The evaluations undertaken for DG programs generally focus on implementation and management concerns and have not collected the data needed for sound impact evaluations. For example, most past evaluations of DG projects have not made comparable baseline and postproject data measurements on key outcomes, and almost all past evaluations lacked data on comparison groups that did not receive assistance. This makes it nearly impossible to develop a retrospective analysis from the data in those evaluations to accurately determine the effects of DG programs.
- There is a tendency, at one and the same time, to evaluate democracy projects mainly in terms of very proximate outcome measures that mainly assess how well the project was implemented and yet to judge the ultimate success of DG projects by whether they coincide with changes in country-level measures of national democracy such as Freedom House scores. Neither course best serves USAID's interests in determining the effects of its DG programs. Those effects are best judged by focusing on policy-relevant objectives at the local or sectoral level that are plausible outcomes of those projects.
- Once research and evaluation are completed, there are few organizational mechanisms for broad discussion among DG officers or for integration of research and evaluation findings with the large range of analysis being carried on outside the agency.
- DG officials are genuinely interested in procedures that will help them better learn and demonstrate the impact of their projects. Yet there is considerable concern among many at USAID regarding whether missions would gain from designing or implementing rigorous impact evaluations, especially those using randomized assignments. This is mostly due to deep skepticism as to the applicability of this methodology to DG programs but also to the overall decline in support for evaluations within USAID, to a lack of specific expertise on impact evaluation design, and to issues in contracting timetables and procedures that discourage adoption of what is perceived as a more complicated approach to evaluation.
- More generally, while there are many calls from policymakers, USAID officials, and other international and national agencies and donors to better determine the effects of DG programs, there is also widespread skepticism regarding whether impact evaluations will, in fact, provide that information. One member of the committee, Larry Garber, emphatically shares these concerns. Among both scholars and policy professionals, skeptics worry that the designs for impact evaluations will prove too cumbersome or inflexible to work in fluid and politically sensitive conditions in the

field; that such evaluations will be too costly or time consuming; or that such studies, in particular randomized designs, are either unethical for or ill suited to the actual projects being carried out in DG programs.

Feasibility of Impact Evaluations for DG Projects

Recognizing the need to take such concerns seriously, the committee examined a wide range of impact evaluation designs and worked with DG officers at several missions to assess the feasibility of such designs for their current or planned activities. The committee's field studies found that a much larger portion of USAID's DG programs than expected—forming roughly half of the projects that examined in Uganda and several projects in Peru and Albania—appear to be amenable, in the view of the committee's consultants, to randomized assignment designs. Nor did these designs necessarily require major departures from current program procedures. Often just more attention to how programs were rolled out or allocated among groups scheduled to receive assistance, combined with measurements on both the groups currently receiving assistance and those scheduled to receive it in the future, would create a reasonable randomized assignment design. In cases where randomized assignment designs were not feasible, the field teams were able to develop other designs that could offer a significant improvement in the ability to assess project effectiveness.

In addition, the committee found that many of the surveys that USAID is already carrying out provide excellent baseline and comparison data for DG projects; thus the data for impact evaluations that use matched or adjusted comparison groups (rather than randomization) are in some cases already being collected and could be utilized for little additional cost.

The field teams thus concluded that it was quite feasible, at least in theory, to conduct high-quality impact evaluations of varied designs that will help USAID better discern the impacts of its DG programs. However, the committee knows that there is much skepticism regarding these procedures, and in particular concerns—noted by Mr. Garber and by others in the democracy assistance donor community—about whether the complexity and sensitivity of DG programs will permit sound impact evaluations, especially those using randomized assignments, to be carried out. Therefore the full committee agreed that the value of such impact evaluations will have to be demonstrated in USAID's own experience.

Strategies for Implementation

- **The committee unanimously recommends that USAID move cautiously but deliberately to implement pilot impact evaluations of several carefully selected projects, including a portion with randomized designs, and expand the use of such impact evaluations as warranted by the results of those pilot evaluations and the needs expressed by USAID mission directors.**
- **Moreover, the committee recommends that these pilot evaluations be undertaken as part of a DG evaluation initiative with senior leadership that will also focus on improving USAID's capacity to undertake impact evaluations and make resources and expertise available to mission directors seeking to learn about and apply impact evaluations to their projects. This DG evaluation initiative is described in more detail below.**

Recommendation 2: Developing Better Sectoral-Level Indicators Measuring Democracy

In response to Charge 2, the committee reviewed the most widely used indicators of a country's overall democratic status and considered a number of alternative approaches to developing an operational definition of democracy. This led to three key findings:

- The concept of democracy cannot, in the present state of scientific knowledge of democracies and democratization, be defined in an authoritative (nonarbitrary) and operational fashion. It is an inherently multidimensional concept, and there is little consensus over its attributes. Definitions range from minimal—a country must choose its leaders through contested elections—to maximal—a country must have universal suffrage, accountable and limited government, sound and fair justice and extensive protection of human rights and political liberties, and economic and social policies that meet popular needs. Moreover, the definition of democracy is itself a moving target; definitions that would have seemed reasonable at one time (such as describing the United States as a democracy in 1900 despite no suffrage for women and major discrimination and little office-holding among minorities) are no longer considered reasonable today.
- Existing empirical indicators of overall democracy in a country suffer from flaws that include problems of definition and aggregation, imprecision, measurement errors, poor data coverage, and a lack of agreement among scales intended to measure the same qualities. There is thus no way to utilize existing macro-level indicators in a way that provides sound policy guidance or reliably tracks modest or short-term changes in a country's democratic status. Existing indicators work best simply to roughly categorize countries as “fully democratic,” “authoritarian,” or “mixed or in between” and to identify large-scale or long-term movements in levels of democracy. They are particularly weak in assessing differences among the nondemocratic and mixed regimes that are the most important settings for USAID's DG work.
- By contrast, indicators focused on specific sectors of democracy in a country (the sectoral level) would help USAID (1) track trends across various dimensions of democracy through time, (2) make precise comparisons across countries and regions, (3) understand the components and possible sequences of democratic transition, (4) analyze causal relationships (e.g., between particular facets of democracy and economic growth), and (5) assess the democratic profile (i.e., strengths and weaknesses across various dimensions of democracy) of countries where USAID operates.
- While the United States, other donor governments, and international agencies that are making policy in the areas of health or economic assistance are able to draw on databases that are compiled and updated at substantial cost by government or multilateral agencies mandated to collect such data, no comparable source of data on democracy at either the macro or sectoral level currently exists. Data on democracy are instead currently compiled by various individual academics on irregular and shoestring budgets, or by NGOs or commercial publishers, using different definitions and indicators of democracy.

Strategies for Implementation

These findings have led the committee to make a recommendation that committee members believe would significantly improve USAID's (and others') ability to track countries' progress and make the type of strategic assessments that will be most helpful for DG programming.

- **USAID and other policymakers should explore making a substantial investment in the systematic collection of democracy indicators at a disaggregated sectoral level—focused on the components of democracy rather than (or in addition to) the overall concept.** If they wish to have access to data on democracy and democratization comparable to the data relied on by policymakers and foreign assistance agencies in the areas of public health or trade and finance, a substantial government or multilateral effort to improve, develop, and maintain international data on levels and detailed aspects of democracy would be needed. This should not only involve multiple agencies and actors in efforts to initially develop a widely accepted set of sectoral data on democracy and democratic development but should also seek to institutionalize the collection and updating of democracy data for a broad clientele, along the lines of the economic, demographic, and trade data collected by the World Bank, the United Nations, and the International Monetary Fund.

While creating better measures at the sectoral level to track democratic change is a long-term process, there is no need to wait on such measures for determining the impact of USAID's DG projects. USAID has already compiled an extensive collection of policy-relevant indicators to track specific changes in government institutions or citizen behavior, such as levels of corruption, levels of participation in local and national decision making, quality of elections, professional level of judges or legislators, or the accountability of the chief executive. **Since these are, in fact, the policy-relevant outcomes that are most plausibly affected by DG projects, the committee recommends that measurement of these factors rather than sectoral-level changes be used to determine whether the projects are having a significant impact on the various elements that compose democratic governance.**

Recommendation 3: Using Case Studies of Democratization and Democracy Assistance

The third charge to the committee was to recommend a plan for comparative historical case studies of DG assistance. A clustered set of case studies, tracing the processes through which advances toward democracy were made from various sets of initial conditions, is an appropriate mode of investigation for these issues. Such case studies could be particularly valuable in mapping out varied trajectories of political development and identifying the role that democracy assistance could play in such trajectories in relation to various actors and events.

Nonetheless, committee members were unable to agree on a firm recommendation that USAID should invest its own funds in such case studies since substantial case study research on democratization is being undertaken by academics and NGOs. To learn more about the role of its DG assistance projects in varied conditions and their role in varied trajectories of democratization, USAID could seek to gain from ongoing academic research. Since much potentially relevant academic research is not written for a policy audience, however, USAID would need to structure its interactions with researchers to ensure that it gains useful and relevant information.

Strategies for Implementation

- **If USAID decides to invest in supporting case study research, the committee recommends using a competitive proposal solicitation process to elicit the best designs. USAID should not specify a precise case study design, but instead should specify key criteria that proposals must meet.** These should include (1) the criteria for choosing cases should be explicit and theoretically driven; (2) the cases should include a variety of initial conditions or contexts in which USAID DG projects operate; (3) the cases should include at least one, if not several, countries in which USAID and other donors have made little or no investment in DG projects; and (4) the cases should include countries with varied outcomes regarding democratic progress or stabilization.
- In addition to case studies, a variety of other research methods, both formal and informal (including debriefings of USAID field officers, statistical analyses of international data, and surveys) can shed light on patterns of democratization as well as how DG projects actually operate in the field and how they are received. **USAID should include these varied sources of information as part of the regular organizational learning activities the committee recommends next.**

Recommendation 4: Rebuilding USAID's Institutional Mechanisms for Learning

Regardless of whether USAID conducts many or fewer impact evaluations and contracts for case studies or works with case studies funded by think tanks or other organizations, little of what is learned will effectively guide and improve DG programming without some mechanism within USAID for learning from its own and others' research on democracy and democratization. For USAID to benefit from this committee's proposed pilot study of impact evaluations, it will need to have regular means of disseminating the results of those and other evaluations throughout the agency and discussing the lessons learned from them. For USAID to benefit from ongoing academic research and the studies of DG assistance being undertaken by think tanks and NGOs, it will be necessary for the agency to organize regular structured interactions between such researchers and its DG staff.

While it will take some time for USAID to learn from undertaking the pilot impact evaluations, it will gain immediately from augmenting its overall learning activities and increasing opportunities for DG staff to actively engage with current research and studies on democratization. Though some committee members believe that the impact evaluations will be more novel and instructive than most current case study and policy reports on democratization, several committee members wish to emphasize the considerable value to policymakers and DG officers of the many books, articles, and reports that have been prepared in recent years by academics, think tanks, and practitioners. Whatever the methodological flaws of these case studies and process evaluations from a rigorous social sciences perspective, the committee notes that this expanding literature has provided important lessons and insights for crafting effective DG programs. Thus the committee is unanimous in finding that a renewed emphasis on engaging USAID DG personnel in discussion and analysis of current research on democratization and democracy assistance—including both varied types of evaluations and a broad range of scholarship—would be worthwhile and should begin even before the pilot evaluations have been completed.

Unfortunately, in recent years USAID has substantially reduced its institutional mechanisms for creating, disseminating, and absorbing knowledge. The Center for

Development Information and Evaluation (CDIE), which served as the hub of systemic evaluation for USAID aid projects, has been dissolved. Moreover, USAID's support of conferences and learning activities for mission directors and DG staff to share experiences and discuss the latest research has declined. And although central collection of evaluations is already a requirement, in practice much useful information, including evaluations and other project documents, survey data and reports, and mission director and DG staff reports, remains dispersed and difficult to access.

Strategies for Implementation

Rebuilding organizational learning capacity within USAID will require a number of steps, some minor and some potentially involving major shifts in organizational procedures. The committee thus recommends that the steps below be undertaken by a special DG evaluation initiative led by a senior policymaker or official within USAID who will have the ability to recommend agency-wide changes, as many of the obstacles to improved learning about DG programs stem from agency-wide procedures and organizational characteristics. While in some ways this will replace the capabilities lost with CDIE, in some ways the committee hopes the new initiative will go beyond that.

The committee's charge is limited to recommendations for improving USAID's ability to evaluate its DG projects, but the committee notes that there could be advantages to making this an agency-wide initiative. USAID implements social programs in many parts of the agency, so the changes the committee recommends could yield much wider benefits.

A DG Evaluation Initiative

In support of Recommendations 1 and 4, the committee recommends that USAID develop a five-year DG evaluation initiative, led by a senior USAID official and with special funding, for the following:

1. Undertaking Pilot Impact Evaluations

The committee strongly recommends that to accelerate the building of a solid core of knowledge regarding project effectiveness, the DG evaluation initiative should immediately develop and undertake a number of well-designed impact evaluations that test the efficacy of key project models or core development hypotheses that guide USAID DG assistance. A portion of these evaluations should use randomized designs, as these are the most accurate and credible means of ascertaining program impact. As randomized designs have also been the most controversial, especially in the DG area, it would be most valuable for the evaluation initiative to help USAID gain experience with and determine the value of these designs for learning the impacts of DG projects.

By key models the committee refers to programs that (1) are implemented in a similar form across multiple countries and (2) receive substantial funding (e.g., local government support, civil society, judicial training). By core hypotheses the committee refers to assumptions guiding USAID program design that, whether drawn from experience or prevailing ideas about how democracy is developed and sustained, have not been tested as empirical propositions.

2. Increasing USAID's Capabilities in Project Evaluation

Supporting the DG evaluation initiative with special, dedicated resources outside the usual project structure would be another signal of a strong commitment to change. It is also important that these resources and accompanying expertise in evaluation design be made available to missions implementing DG programs, so that more rigorous evaluations become an opportunity for missions to gain support, rather than an additional unfunded burden. Any changes to M&E of DG programs will be carried out in the field by over 100 missions and hundreds of implementing partners. Even with the centralization of program and budget decision making undertaken in the foreign assistance reforms of 2006, USAID is a highly decentralized agency and mission staff have substantial discretion in how they implement and manage their programs. **The initiative should thus make its resources and expertise available to mission directors who want its support in conducting impact evaluations or otherwise changing their mix of M&E activities, in order to make the initiative an asset to the DG directors in the field rather than an additional unfunded burden.**

3. Providing Technical Expertise

In recent years as USAID has reduced the number of evaluations it conducts, the agency has also failed to hire experts in the latest evaluation practices to guide and oversee its contracting and research. **The committee recommends that USAID acquire sufficient internal expertise in this area to both guide an initiative at USAID headquarters and provide advice and support to field missions, as a key element of the initiative.**

4. Improving the Ease of Undertaking Impact Evaluations of DG projects

While many evaluations are currently only sought well after a project has begun or even only after its completion, impact evaluations generally require before-and-after measures and data from comparison or control groups that should be designed into the program from its inception and often cannot be obtained at all once a program is well under way. Pressures to get projects under way, as well as many current contracting practices, thus work against implementing and sustaining impact evaluation designs. **One task of the DG evaluation initiative should be to address these issues and explore how to ease the task of undertaking impact evaluations within USAID's contracting and program procedures. The initiative should also examine incentives for both DG officers and project implementers to carry out sound impact evaluations of selected DG projects.**

5. Consider Creating a Social Sciences Advisory Group

To assist in the evaluation effort, the committee recommends that the administrator consider establishing a social sciences advisory group for USAID. This group could play a useful role in advising on the design of the DG evaluation initiative, helping to work through issues that arise during implementation, and developing a peer review process for assessing the evaluations undertaken during the initiative.

6. Rebuilding Institutional Learning Capacity

This initiative should be guided by a policy statement outlining the strategic role of developing USAID as a learning organization in the democracy sector. The committee believes that increasing USAID's capacity to learn what works and what does not should include provisions for regular face-to-face

interactions among DG officers, implementers, and outside experts to discuss recent findings, both from the agency's own evaluations of all kinds and studies by other donors, think tanks, and academics. Videoconferencing and other advanced technologies can be an important supplement, but personal contact and discussion would be extremely important to sharing experiences of success and failure as the evaluation initiative went forward. This includes lessons about both the effectiveness of DG projects and successes and failures in implementing impact evaluations.

Such meetings are especially important for ensuring that the varied insights derived from impact and process evaluations, academic studies, and examinations of democracy assistance undertaken by independent researchers, NGOs, think tanks, and other donors are absorbed, discussed, and drawn into USAID DG planning and implementation. While only USAID has the ability to develop and carry out rigorous evaluations of its projects' impacts, many organizations are carrying out studies of various aspects of democracy assistance, and USAID's staff can benefit from the wide range of insights, hypotheses, and "lessons learned" that are being generated by the broader community involved with democracy promotion.

Results of the Initiative

At the end of this five-year period, USAID would have:

- Practical experience in implementing impact evaluation designs that will indicate where such approaches are feasible, what the major obstacles are to wider implementation, and whether and how these obstacles can be overcome.
- Where the evaluations prove feasible, a solid empirical foundation for assessing the validity of some of the key assumptions that underlie DG projects and rigorous determinations of the impact of commonly used DG projects in achieving program goals.
- A core of expertise within USAID on the latest evaluation methods and practices.
- Institutionalized learning practices across the organization to keep officials engaged, informed, and up-to-date on the latest findings from within and outside USAID regarding democracy and democracy assistance.

Conclusion

The committee stresses that the goal of USAID should not be merely incremental improvement of its project evaluations, or funding additional case studies, but building the entire capacity of the agency to generate, absorb, and disseminate knowledge regarding democracy assistance and its effects. This will necessarily involve (1) gaining experience with varied impact evaluation designs, including randomized studies, to ascertain how useful they could be for determining the effects of DG projects; (2) focusing on disaggregated, sectoral-level measures to track democratic change; (3) expanding the diversity of case studies that are used to inform thinking on DG planning; and (4) adopting mechanisms and activities to support the active engagement of DG staff and mission personnel with new research on democratization and DG assistance.

REFERENCES

Finkel, S.E., Pérez-Liñán, A., and Seligson, M.A. 2007. The Effects of U.S. Foreign Assistance on Democracy Building, 1990-2003. *World Politics* 59(3):404-439.

Finkel, S.E., Pérez-Liñán, A., Seligson, M.A., and Tate, C.N. 2008. Deepening Our Understanding of the Effects of U.S. Foreign Assistance on Democracy Building: Final Report. Available at: <http://www.LapopSurveys.org>.

IMPROVING DEMOCRACY ASSISTANCE: Building Knowledge Through Evaluations and Research

**Committee on Evaluation of
USAID Democracy Assistance Programs**

**Development, Security, and Cooperation
Policy and Global Affairs**

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

PREPUBLICATION COPY
UNCORRECTED PROOFS

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, N.W. Washington, DC 20001

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This study was supported by Contract No. DFD-C-00-06-00091-0 between the National Academy of Sciences and the U.S. Agency for International Development. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the organizations or agencies that provided support for the project.

International Standard Book Number xxxxxxxxxxxx
Library of Congress Catalog Card Number xxxxxxxxxxxx

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, N.W., Lockbox 285, Washington, DC 20055; (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area); Internet, <http://www.nap.edu>.

Copyright 2008 by the National Academy of Sciences. All rights reserved.

Printed in the United States of America

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

Committee on the Evaluation of USAID Democracy Assistance Programs

Jack A. Goldstone (*Chair*), George Mason University, Fairfax, VA

Larry Garber, New Israel Fund, Washington, DC

John Gerring, Boston University, Boston, MA

Clark C. Gibson, University of California, San Diego, La Jolla, CA

Mitchell A. Seligson, Vanderbilt University, Nashville, TN

Jeremy Weinstein, Stanford University, Stanford, CA

Staff:

Jo L. Husbands, Senior Project Director

Paul Stern, Board Director

Tabitha Benney, Senior Program Associate*

Rita Guenther, Senior Program Associate

*Until August 2007.

PREFACE

Since September 11, 2001, the Bush administration has made support for democracy one of the major pillars of U.S. security policy. The U.S. Agency for International Development (USAID) has been providing democracy assistance to countries around the world for over 25 years and has invested substantially in a variety of programs in diverse political situations. To better understand the impact of its democracy assistance efforts, the agency launched the Strategic and Operational Research Agenda (SORA). As part of SORA's work, USAID asked the National Research Council to prepare a report on how best to evaluate USAID democracy and governance (DG) programs.

The National Academies appointed an ad hoc committee to work on this report, including scholars with long experience and varied methodological approaches to the study of democracy and democratization, and a former USAID mission director with field experience in implementing DG programs. I extend my deepest personal thanks to each of them for their many intellectual contributions to the committee's work and for the time and effort they gave to the report. It was a pleasure to work with such outstanding colleagues.

To fulfill the mission given to the Academies, additional scholars were called on to help the committee examine key methodological issues in evaluating the impact of DG assistance. The committee's deliberations with these scholars included a conference in Boston, Massachusetts, on issues in measuring democracy and a conference in Stanford, California, on how case studies of democratization and democracy assistance could inform DG programming. The committee also contracted with several expert consultants in the design and implementation of program evaluations, from both the academic and policy implementation spheres, to visit several USAID missions to examine the feasibility and scope for developing impact evaluations of DG projects in the field. The committee owes great thanks to these scholars for their contributions to this report. (A full listing of participants and consultants is given in Appendixes A and E).

The committee spent many sessions discussing evaluation procedures with representatives of USAID and former and current contractors for the agency's DG programs. The committee is grateful to USAID for providing the time and assistance to set up these meetings and for their willingness to work through ideas and opportunities with the committee. In particular, the committee thanks three USAID officials who were the primary contacts throughout the project and whose support and advice were critical to the success of the committee's work: Margaret Sarles, chief of the Strategic Planning and Research Division, and two members of the SORA staff, David Black, who served as project officer, and Mark Billera. David and Mark also accompanied the committee's teams on their field visits. The goal was not merely to recommend abstract "ivory tower" ideas regarding project evaluations but to learn from USAID and provide recommendations that would be feasible in the field and useful on a variety of levels for USAID planning and program implementation.

The committee particularly wants to thank the USAID missions in Albania, Peru, and Uganda, who hosted committee members, staff, and consultants, and the USAID Washington officers who helped arrange those visits. The field visits were invaluable in determining how actual DG programs were being evaluated and learning how USAID staff and consultants could develop different evaluation designs for current and forthcoming USAID programs.

The committee also thanks the members of the National Research Council staff who provided substantive and administrative support for the project. Jo Husbands served as project director and helped guide the committee through dozens of meetings, lengthy deliberations, and the administrative hurdles of carrying out the committee's ambitious goals. She also made substantial contributions to the many drafts of the report. Paul Stern offered sage advice throughout the process, particularly on methodological and measurement issues. Rita Guenther and Tabitha Benney provided research and administrative support, along with immense energy and good cheer. Rita also took the lead in drafting the report summarizing the three field visits and drafted several sections of the report. Three of them also took part in the field visits—Jo in Albania and Uganda, Rita in Albania and Peru, and Tabitha in Peru.

At a time when democracy assistance is becoming ever more important as part of international and U.S. policies to assist developing nations, build peace, and reduce conflict, the committee hopes this report can serve as a practical guide to policymakers and USAID mission staff. Foreign assistance donors and aid organizations in a variety of areas are demanding better proof of results and more certain knowledge on which to build future assistance programs. The committee provides recommendations on how USAID can design its activities to gain greater knowledge of which DG projects are most effective in the field and how to use that knowledge—drawing on both internal experience and outside expertise—to guide and improve future democracy assistance. It is hoped that the recommendations in this report will lead to not only more effective programs to assist the emergence and stabilization of democracies but also the adoption of evaluation methods that will improve aid effectiveness throughout the domain of U.S. foreign assistance.

Jack A. Goldstone

Chair, Committee on the Evaluation of USAID Democracy Assistance Programs

Acknowledgments

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the National Research Council's Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process. We wish to thank the following individuals for their review of this report: Valerie Bunce, Cornell University; Kenneth Bollen, University of North Carolina, Chapel Hill; Susan Hyde, Yale University; Robert Keohane, Princeton University; David Laitin, Stanford University; Carol Lancaster, Georgetown University; Ruth Levine, Center for Global Development; Michael Lund, Management Systems International; Gerald Munck, University of Southern California; Barbara Torrey, Population Reference Bureau, Inc.; and Nicholas van de Walle, Cornell University.

Although the reviewers listed above provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations, nor did they see the final draft of the report before its release. The review of this report was overseen by Charles Tilly, Columbia University, and Enriqueta C. Bond, Burroughs Wellcome Fund. Appointed by the National Research Council, they were responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring committee and the institution.

Contents

Summary	1
Chapter 1: Democracy Assistance and USAID	15
U.S. Democracy Assistance: A Brief Introduction	15
Democratic Development and Democracy Assistance: What Do We Know? 19	
USAID's Request to the National Research Council	22
Report Overview	25
References	32
Chapter 2: Evaluation in USAID DG Programs: Current Practices and Problems	39
Introduction	39
Current Evaluation Practices in Development Assistance: General Observations	39
Current Policy and Legal Framework for USAID DG Assessments and Evaluations	46
Three Key Problems with Current USAID Monitoring and Evaluation Practices	49
Conclusions	55
References	58
Chapter 3: Measuring Democracy	61
Introduction	61
Problems with Extant Indicators	63
A Disaggregated Approach to Measurement at the Country Level	70
Conclusions	78
References	80
Chapter 4: Learning from the Past: Using Case Studies of Democratic Transitions to Inform Democracy Assistance	85
Introduction	85
Case Study Designs and Methods	85
Insights from Current Research: Results of a Conference of Case Study Specialists on Democracy	87
A Multicase Study Design to Generate and Investigate Strategic Hypotheses Regarding Democracy Assistance	94
Conclusions	98
References	99
Chapter 5: Methodologies of Impact Evaluation	103
Introduction	103
Importance of Sound and Credible Impact Evaluations for DG Assistance	103
Plan of This Chapter	106
Points of Clarification	107

Internal Validity, External Validity, and Building Knowledge	109
A Typology of Impact Evaluation Design	113
Examples of the Use of Randomized Evaluations in Impact Evaluations of Development Assistance (Including DG Projects)	122
References	126
Chapter 6: Implementing Impact Evaluations in the Field	131
Introduction	131
Field Visits to USAID Missions	131
Employing Randomized Impact Evaluations for USAID DG Projects in the Field	132
Challenges in Applying Randomized Evaluation to DG Programs	143
Conclusions	148
References	150
Chapter 7: Additional Impact Evaluation Designs and Essential Tools for Better Project Evaluations	151
How Often Are Randomized Evaluations Feasible?	151
Designing Impact Evaluations When Randomization Is Not Possible	153
What to Do When There Is Only One Unit of Analysis?	161
Conclusions	164
References	166
Chapter 8: Creating the Conditions for Conducting High-Quality Evaluations of Democracy Assistance Programs and Enhancing Organizational Learning	167
Introduction	167
Issues in Obtaining High-Quality Impact Evaluations	167
Improving Organizational Learning	173
Conclusions	180
References	180
Chapter 9: An Evaluation Initiative to Support Learning the Impact of USAID's DG Programs	181
Introduction	181
Providing Leadership and Strategic Vision	182
Implementing the Vision: The Evaluation Initiative	184
Agenda for USAID and SORA	187
Role of Congress and The Executive Branch	189
Conclusions	190
References	191
Glossary	193
Appendix A: Biographical Sketches of Committee Members	199
Appendix B: Committee Meetings and Participants	202
Appendix C: Measuring Democracy	215
Appendix D: Understanding Democratic Transitions and Consolidations from Case Studies: Lessons for Democracy Assistance	245

Appendix E: Field Visit Summary Report 250

Appendix F: Voices from the Field: Model Questionnaire 272

