

Electing to Evaluate: A Review of *Evaluating Democracy Assistance*

Evaluating Democracy Assistance

Krishna Kumar

(Lynne Rienner Publishers, 241 pp., \$25.00)

By Jerome Wei

Over forty years ago, foundations run by West Germany's political parties successfully helped Spain and Portugal's transition to democracy by providing training and funding for democratic groups in those countries (Pinto-Duschinsky 1991). The National Endowment for Democracy, the United States' first organization dedicated to supporting democracy around the world, was founded in 1983 and was designed to emulate the German foundations' model. The field of democracy assistance, where international actors seek to develop and deepen democracy in other countries, was born. Since then, promoting democracy has become a key component of the United States' and Europe's foreign policies and funding for democracy assistance has steadily increased. In 2014, the United States alone allocated over \$2.8 billion to support democracy, good governance, and human rights overseas.

Despite decades of concerted effort by democracy promoters, prominent scholars and essayists argue that the world today is in a democratic recession (Plattner 2015). The failures of the Arab Spring in 2011 have led

to instability and war, along with the collapse of state institutions, in the Middle East and North Africa. In West Africa, weak democratic institutions are hard-pressed to deal with health crises and Islamic radicalism. Russia, China, and other authoritarian states are increasing their repression of civil society and the media. Whether the world is actually experiencing a marked democratic decline or not (and there remains a significant debate about this), many in the democracy assistance community have taken a step back, reappraising their strategies and the global operating environment for their work. The effective evaluation of democracy assistance has never been needed more than it is today.

Though the strategies and methods of democracy promoters have been studied extensively, little has been written on evaluating democracy assistance. Krishna Kumar, in *Evaluating Democracy Assistance*, seeks to fill this gap by overviewing the evaluation landscape for democracy assistance. The Senior Evaluation Adviser in the State Department's Office of Foreign Assistance Resources

and formerly at the US Agency for International Development's (USAID) Center for Development Information and Evaluation, Kumar offers valuable advice to evaluators, program designers, and program implementers on how to design and manage evaluations of democracy assistance. In the eight chapters of his book, Kumar systematically covers the basic steps of program evaluation in a democracy assistance context—indicators, monitoring, evaluation planning, evaluation design, data collection, and communication of evaluation findings—while discussing tradeoffs, best practices, and improvements along the way.

However, this is not just a valuable operational handbook for evaluators. Kumar also advances the important argument that core characteristics of democracy assistance pose challenges for those seeking to apply evaluation techniques from the development realm to democracy assistance. By doing so, he continues the difficult work of academics and practitioners who have attempted to draw a conceptual line between democracy assistance and traditional development programs (for more on this topic, see Carothers and de Gramont 2013).

In chapter 2, Kumar breaks down these challenges into three obstacles. First, the underlying intervention models for democracy programs lack both theoretical and empirical support. Kumar notes just how little consensus there is in political science over the many subject areas (electoral systems, corruption, political party development) that comprise democracy assistance. The lack of academic consensus means that many programs do not operate off an underlying intervention model (often called a theory of change) that is sound and verifiable. An evaluator

looking at the impact of a program, then, will not only look at the outputs and outcomes of a program in relation to its model but may also have to reexamine the model itself. As Kumar rightly points out, time and resource constraints will often limit the ability of evaluators to conduct an evaluation that addresses both impact and the validity of the model. Also important but unmentioned by Kumar is the fact that many evaluators, especially those from outside academia, lack the expertise and background to analyze and propose alternate models for democracy assistance. USAID, seeking to remedy this analytical gap, began a project in 2014 with academic researchers to identify theory of change models that will generate stronger designs and evaluations of democracy programs (IIE 2015).

Kumar's second obstacle is compelling but also presents a rather intractable problem: the large number of factors affecting democratic development and behavior makes it difficult to isolate and attribute impacts to any one democracy program. While it can be easy to measure outputs and even short-term outcomes, democracy assistance seeks to change long-term behavior and practices in institutions and society. We cannot assume that training members of political parties will lead to less political corruption over the long run because a range of other factors—education, income, social norms, policing, media—all affect the honesty of politicians. There is no easy fix for the attribution problem. Kumar believes that only long-term academic research can answer fundamental questions about impacts and the mechanisms of democratic change; his call for closer ties between academics and practitioners is a consistent theme throughout the book and a valuable recommendation.

In many countries, the politi-

cal environment is not conducive for data collection in the field (imagine, for example, the difficulty of surveying citizen perceptions of democracy in closed societies like Eritrea or Belarus). Difficulty collecting data is Kumar's last obstacle and one of vast practical significance. Democracy assistance often operates where democracy is the weakest; these are the same places where evaluations are hardest to conduct. All three of the aforementioned obstacles infuse the rest of his book; he describes the impacts of these obstacles at each step of the evaluation process and details strategies for mitigating them.

Indicators and rankings provide evaluators with ready-made data, a potential solution to the obstacle of collecting field data in a hostile political environment. Kumar devotes chapter 3 to these democracy indicators, giving an overview and reliability assessment of macro-level indices and meso-level indices. Macro-level indices measure the state of democracy at the country level while meso-level indices measure the strength of democratic sectors such as media, civil society, human rights, and elections. Kumar makes a valid point in this section: while macro-level indicators are useful in newspaper reporting and academic studies of global change in democracy, they are of limited use for evaluators analyzing democracy assistance programs at the national or sub-national level. These indicators generally lack the sensitivity to track incremental changes in the quality of democracy over time or across countries. What is more important, however, is the attention Kumar pays to meso-level indicators, which are much more useful for practitioners. These measures home in on specific sectors, overlapping with sector-building programs that are common in democracy assistance.

While Kumar notes that meso-level indicators also lack sensitivity to incremental changes over time within a sector, these sector indicators can be improved and made more rigorous through evaluation experience and collaboration with university researchers.

Many readers will be drawn to Kumar's section on experimental and quasi-experimental designs in chapter 6. Experimental designs construct treatment and control groups through random sampling while quasi-experimental designs use equivalent comparison groups without random sampling. These designs, most prominently the randomized control trial (RCT), isolate the effects of a program from the effects caused by external factors and are becoming increasingly popular in the development world. Proponents believe that these strategies will mitigate the effects of Kumar's second obstacle, allowing evaluators to attribute impacts to democracy assistance programs with greater confidence. Kumar, however, approaches these new designs more cautiously. He identifies 11 methodological and practical obstacles that organizations face when using these evaluation techniques.

Several of these 11 challenges require additional discussion. The first is the fact that many democracy assistance programs seek to make changes at the national level: influencing or improving national political parties, policies, or institutions. The obstacle in this context is that control or comparison groups cannot be constructed for these programs, precluding the use of experimental and quasi-experimental designs. The targeting of programs at the national level will continue to be the most important factor limiting the use of experimental designs in the field of democracy assistance. Other types of democ-

racy assistance programs promote knowledge and civic education of the population or within civil society organizations. Programs promoting knowledge and civic education present the second challenge highlighted here. Evaluators using experimental designs to assess the impact of these programs have to overcome “contamination effects”—the simple fact that people and organizations share information outside the context of a program, reducing the comparability of a treatment and comparison group. Finally, Kumar notes that many programs have not been designed with experimental or quasi-experimental designs in mind, as they do not collect baseline data. While Kumar explains that in many cases there is a simple lack of technical expertise to plan and manage baseline data collection, he does not take the opportunity to push democracy assistance organizations to incorporate baseline data collection and experimental design-friendly features into future programs where possible.

Kumar believes that experimental and quasi-experimental designs are of limited usefulness for democracy assistance. Even so, he finds that there are three areas of value for these strategies: to measure the impact of focused, short-term interventions; to compare the effectiveness of different program delivery methods; and to test the hypotheses and models underlying democracy interventions. Kumar is too pessimistic on the applicability of experimental and quasi-experimental designs in democracy assistance. While Kumar does not claim to be comprehensive in his book, his analysis does not incorporate the experience of organizations like the MIT-based research organization J-PAL, which has made its name popularizing RCTs. J-PAL has conducted over 100 evaluations

of programs classified as governance programs using experimental and quasi-experimental designs (J-PAL 2015). These evaluations surely hold many lessons for the democracy assistance community. Furthermore, as mentioned in the previous paragraph, these techniques would be more useful if more baseline data collection were incorporated into program design. Despite these critiques, Kumar’s central contribution in this section is the much-needed context he provides to the use of experimental and quasi-experimental designs in democracy assistance. This section of the book is a reminder to stakeholders who span the development and democracy assistance worlds that there are important differences between the two.

Non-experimental designs do not use a control or comparison group and can be statistical (pre- and posttest and cross-sectional designs) or case study-based. Although experimental designs are increasingly popular, non-experimental designs, especially case studies, continue to be the most common evaluation format in democracy assistance. While finding case studies the most suitable evaluation design for democracy evaluations, Kumar finds that there are plenty of areas in which evaluators can improve their methodological rigor. He notes that because there are oftentimes no clear rules for identifying a finding, no verification of findings against other data, and no clear frameworks for analysis, case study-based evaluation reports may look imprecise and give the impression of bias.

In chapter 7, Kumar makes five suggestions to improve case studies. First, evaluators should focus on the evaluation questions outlined in their terms of reference or statements of work. Second, their evaluations should use counterfactual reasoning;

that is, what would have happened in the absence of the program. Counterfactual reasoning can more clearly outline impacts in a case study-based evaluation and is currently lacking. Kumar's third suggestion is that evaluations enlist more host-country researchers. Host-country researchers elicit more balanced and honest responses during data collection and know more about the local context than Western evaluators do. The fourth suggestion is for evaluators to pay more attention to sampling when interviewing key informants or designing surveys. He finds that most evaluations to date have relied on nonprobability sampling methods, which often result in unrepresentative samples and thus limit the validity of the collected data. Finally, Kumar encourages greater use of quantitative data in case study designs, primarily from surveys, structured observations, and secondary data such as the findings of other surveys, opinion polls, and other records.

Much of the discussion in the democracy assistance community surrounding Kumar's book is concerned with the author's stance on experimental and quasi-experimental designs. However, the central contri-

bution of this book for practitioners is that it highlights gaps and deficiencies in current practices. The section on case study designs is important because he highlights these gaps and offers simple and achievable fixes that may significantly improve the reliability of the majority of evaluations done in democracy assistance.

Kumar's book is a useful tool for evaluators working in the democracy assistance world, managers looking to improve their democracy interventions, and interested readers looking for an introduction to program evaluation through a distinctive field in the world of foreign assistance. He outlines current gaps in the field of democracy evaluations and offers valuable suggestions for improvement. Kumar lays the ground for future research and a roadmap for the democracy community to work together to improve our understanding of how democratic change comes about and how outside actors can facilitate that change. Though an important contribution, Kumar's is not the final word on evaluations of democracy assistance; I hope his book will spur other practitioners and academics to contribute to the field.

References

- Carothers, Thomas, and Diane de Gramont. 2013. *Development Aid Confronts Politics: The Almost Revolution*. Washington, DC: Carnegie Endowment for International Peace.
- IIE (Institute of International Education). 2015. "Grant: Theories of Change." Accessed April 7. <http://www.iie.org/en/Programs/USAID-Democracy-Fellows-and-Grants-Program/G-Theories-of-Change>.
- J-PAL (Abdul Latif Jameel Poverty Action Lab). 2015. *Evaluations*. Accessed April 7. <http://www.povertyactionlab.org/GI/Evaluations>.
- Pinto-Duschinsky, Michael. 1991. "Foreign Political Aid: The German Political Foundations and Their US Counterparts." *International Affairs* 67(1):33–63.
- Plattner, Marc F. 2015. "Is Democracy in Decline?" *Journal of Democracy* 26(1):5–10.

Jerome Wei, who also goes by Jerry, is a first-year Master of Public Policy candidate at the George Washington University concentrating in program evaluation and quantitative methods. He hopes to foster better institutional learning, planning, and oversight in federal agencies. He graduated from Yale University in 2010 with a Bachelor of Arts in Political Science. After graduating, he worked for four years as a Program Planning Analyst at the National Endowment for Democracy, a foundation that supports democracy and human rights around the world. He enjoys exploring the DC suburbs for Southeast Asian cuisine, going to area book sales, and watching gloomy drama series on HBO.

The reviewer would like to thank Editor-in-Chief Anne Kruse and especially Senior Associate Editor Odia B. Cissé for the time they spent making this book review accurate, coherent, and, frankly, structured like a book review. The reviewer would also like to thank his mentors at Yale, Susan Hyde and Nikolay Marinov, who introduced him to democracy assistance, and his bosses at the National Endowment for Democracy, Georges Fauriol and Myriam Fizazi-Hawkins, who gave him unparalleled opportunities to pursue his professional interests.