

**The Devolution Initiative Evaluation:
Innovation and Learning at the W. K. Kellogg Foundation**

By Kay E. Sherwood

Edited by Patricia Patrizi and Michael Quinn Patton

This case was developed as part of the Evaluation Roundtable, an association of evaluation professionals working in philanthropy dedicated to developing a community of practice and learning for their work. The development of this case and others has been supported financially by the California Endowment, the Edna McConnell Clark Foundation, the Robert Wood Johnson Foundation, the W. K. Kellogg Foundation, the John S. and James L. Knight Foundation, the David and Lucile Packard Foundation, the James Irvine Foundation, and the Wallace Foundation. Teaching of the cases has also been supported by the Evaluation Roundtable and Grantmakers for Effective Organizations, as well as philanthropic supporters of particular case teaching sessions.

In the fall of 1996, government in the United States embarked on a journey called “devolution” – shifting powers, responsibilities, and funding from the federal level of government to the state level, and sometimes to the local level, for a number of social welfare programs, beginning with the cash assistance program for low-income families formerly known as Aid to Families with Dependent Children (AFDC), often called “welfare.” The embarkation event, a law called the Personal Responsibility and Work Opportunities Reconciliation Act (PRWORA), was radical, overturning a social safety net that had been built over several decades. And it was politically divisive because the shift of authority away from the federal government was accompanied by a shift from an open-ended system of income assistance to one with budget limits, time limits for receiving cash assistance, and requirements for cash assistance recipients to work. Advocates for low-income families worried that many would be left destitute and their children consigned to foster care; public policy analysts worried that states, now freer to set welfare policies, would begin “a race to the bottom” to reduce benefits and assistance to low-income families; the authors of the PRWORA legislation worried that Congressional intent would be undermined by liberals in government and in nonprofit entities seeking to “soften” the provisions of the law.¹

Looking back in 2002 on such concerns, President William Richardson² of the W. K. Kellogg Foundation said,

Even as perhaps the biggest economic boom in history swept the nation, we wondered how Temporary Assistance to Needy Families [which replaced AFDC] would fare when the economy faltered. Even as stocks rose to record highs, we asked what would happen if lower tax revenues required states to cut transportation and childcare supports to the working poor. Even when some financial advisors thought a bear market was unlikely, we questioned how the new safety net programs would function during a lukewarm stage of the business cycle.³

The W. K. Kellogg Foundation had historically focused on communities and community-level issues, and had historically been disinclined to engage public policy issues. The Foundation was also historically reluctant to support research, including policy research. But, as national welfare reform and devolution were being discussed, a convergence of events and changes in the Foundation led to an interest in devolution and related research.

¹ In addition to cash assistance for low-income families, “devolved” by the 1996 law, the Balanced Budget Act of 1997 authorized federally funded health insurance coverage for children that was to be designed and managed by states; this program, called the State Children’s Health Insurance Program (SCHIP), became another key subject of devolution and of the W. K. Kellogg Foundation’s Devolution Initiative.

² William C. Richardson retired as President of the W. K. Kellogg Foundation in December, 2005.

³ William C. Richardson, “Devolution: The End of the Beginning” (speech to Devolution Initiative conference), January 10, 2002, Washington Marriott Hotel, Washington, DC.

First, William (Bill) Richardson had been hired as the new foundation president in 1995. As a scholar of health care policy, Board Chair of the Kaiser Family Foundation, and former President of Johns Hopkins University, he had a longstanding interest in public policy. Richardson also had a three-decade relationship with the Foundation, first as a Kellogg Fellow at the University of Chicago in the 1960's, then as a member of several Foundation task forces, and as a grantee for three projects. Richardson “knew that community-based work and policy research could be combined” at the Foundation. He was also specifically interested in devolution because it was big and had implications for communities.

Second, several leading national research institutions were shopping proposals to funders to gather resources to study the effects of devolution. Richardson saw these as a way to get involved in devolution and public policy if they could be linked with building community capacity in terms of policy change. As he characterized his early idea: “Here is something important about to happen [devolution]; here are two or three organizations coming to us; let's see if it can come together.”

Third, Richardson had a vision for the Foundation to move it “from grantmaking to changemaking” that entailed organizational changes. He initiated a restructuring of the Foundation intended to eliminate independent fiefdoms, increase coordination across programs, and increase the impact of the Foundation's programs – which ultimately provided the form for the Foundation's first large, cross-foundation initiative. In 1996, he created a cross-foundation task force to look at what the Foundation might do to highlight the devolution-related policy changes for the communities historically of interest to the Foundation. The result was the Devolution Initiative, a seven-year, \$56 million project with 31 grantees and a \$3.6 million external evaluation.

Initiative and Evaluation: Co-Evolution

The Devolution Initiative (DI) and its evaluation intentionally unfolded together. The devolution policy was seen as extremely complex, but not predictable in 1996, and the Foundation wanted an evaluation that would help guide its work on the Initiative.

The Initiative: A Focus on Informing Policy

The Foundation staff began the Initiative with an information-focused concept that tied together the historic Foundation mission of “helping people help themselves”⁴ and the emerging policy landscape of devolution. The staff assumed that actors in state, local, and tribal arenas would need information to help them make and influence the making of good policy in the circumstances of devolution. Thus, producing, disseminating, and building capacity to effectively use information was at the heart of the Foundation's Devolution Initiative. The 31 DI grantees played various roles in its

⁴ The fuller version of the W. K. Kellogg Foundation's mission is “to help people help themselves through the practical application of knowledge and resources to improve their quality of life and that of future generations.”

information-based strategy to shape policy. “Partners,” as the grantees were called, included:

- National nonprofit research institutions conducting national research to study the effects of devolution on low-income families. These groups included the Urban Institute, the Nelson A. Rockefeller Institute of Government at the State University of New York-Albany, MDRC, the Hudson Institute, and Johns Hopkins University’s Department of Sociology.
- National policy research organizations tracking and interpreting changes in laws, regulations, funding, programs, demographics, and other trends related to devolution. These groups included the Center on Budget and Policy Priorities, the Center for Law and Social Policy, the Center for Community Change, the Center for Policy Alternatives, the Joint Center for Political and Economic Studies, the National Conference of State Legislatures, and the National Governors’ Association.
- National advocacy organizations – some with state affiliates – tracking the potential effects of devolution on particular populations or existing service programs and working to minimize harms and maximize benefits for these populations and service programs as states, localities, and tribes made decisions about welfare, health care and other social safety net issues. These groups included Families USA, the National Association of Child Advocates (now called Voices for America’s Children), the National Coalition on Health Care, the National Congress of American Indians, the Immigrant Welfare Implementation Collaborative, and the Children’s Defense Fund.
- State advocacy groups – some affiliated with national advocacy organizations – in Florida, Mississippi, Montana, New York, Washington, and Wisconsin working to shape state, local, and tribal social safety net policies.

The Foundation staff also saw in devolution new opportunities for citizen participation in government; in line with the historic mission, it was assumed that the people who would be affected by devolution of social and health care policies should be encouraged and empowered to participate in the state, local, tribal, and national decisions affecting them, their families, and their communities. While citizen mobilization and coalition building were envisioned as key activities of the Initiative from the outset, it was not originally envisioned that the Foundation would fund state-level advocacy groups to do this work. Initially, the assumption was that by funding national advocacy organizations and linking them to national research and policy analysis groups, the DI would generate policy-oriented activity in 40 states because the national advocacy grantees had affiliates or related entities in that many states. (In the Initiative, the Foundation was notably *not* promoting any particular policy, taking care to avoid prohibited lobbying activity.)

The DI goals, as set forth in a progress report to the Foundation's Board in 1999 were stated as:

- Creating an objective information base about the impact of devolution.
- Sharing the findings with policymakers and the public.
- Using the information and other community resources to promote public participation in informing policy agenda and decisions.⁵

The Foundation's early decisions about the Initiative arose from two sources. The first was a brief "premises" paper that laid the foundations for the Initiative and its evaluation:⁶ The premises described were:

Premise 1: While devolution creates new potential for citizen involvement, participation in devolution processes can be enhanced by increasing the capacity and confidence of grassroots and local leaders to influence policy. Capacity building is an activity that can be effectively carried out by intermediary organizations with statewide coalitions.⁷

Premise 2: It is possible to build essential learning relationships among a set of grantees. Through the Devolution Strategy, the grantees will produce, enhance, and capture information from their constituencies and target audiences and get information to them in a two-way facilitated exchange of information. Innovative means and creative channels of dissemination will need to be pursued which facilitate moving relevant information to the right people at the right time. The grantees will build individual and collective knowledge which will provide ongoing influence in decision making processes regarding devolution, and will reach and involve other groups in their learning process. However, we recognize that information packaging, dissemination, and utilization are necessary – but by themselves insufficient – to galvanize participation in policy processes.

Premise 3: The Devolution Strategy represents an innovative form of grantmaking for WKKF⁸ because it combines in a single strategy information producers, advocacy organizations, a policy focus, and the involvement of program directors from multiple programming areas across the Foundation. We want to learn the extent to which grantmaking can create interrelationships among nontraditional partner grantees as a result of our

⁵ "The Devolution Initiative: Application of Knowledge to the Problems of People," A Program Initiative Progress Report, April 14-15, 1999. This formulation of the goals of the Initiative, or close approximations of it, appears in most of the Foundation's documentation and on its website.

⁶ W. K. Kellogg Foundation, "Devolution Initiative Premise Paper," December 18, 1996.

⁷ This meant *national* intermediary organizations with state affiliates – a premise that would be revisited in 1999.

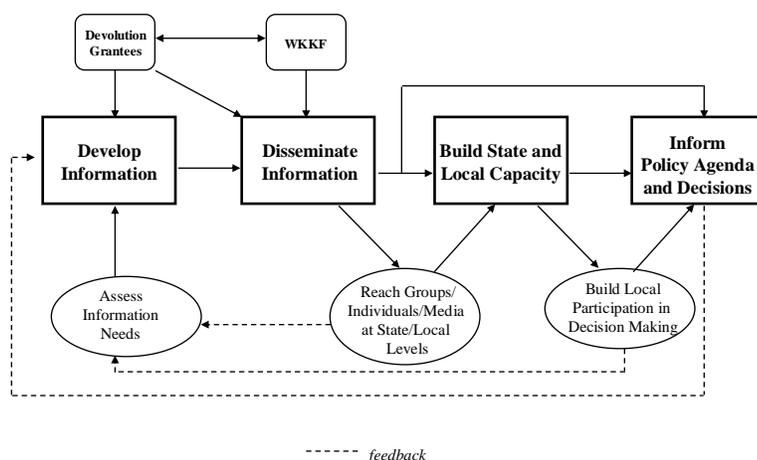
⁸ WKKF is the acronym used internally and by the Devolution Initiative evaluators to refer to the Kellogg Foundation.

investment that go beyond the activities of individual projects. In other words, we intend to determine if it is possible to create working conditions among the grantees that will yield policy effects such that the sum is greater than individual contributions.

Premise 4: Devolution is likely to cause significant changes in relationships between nonprofit organizations, government at all levels, and philanthropy. The Devolution Strategy is an opportunity for the WKKF to participate in a co-evolutionary process with its five grantee partners⁹ that will ultimately allow each participant to continue to be an effective player in a devolved world. WKKF’s funding presence in devolution creates unique opportunities for organizational learning that will allow WKKF to be responsive to a changing environment while at the same time remaining true to its mission of helping people help themselves.

This premises document was authored by Ricardo Millet, the Foundation’s evaluation director, Mark Lelle, an evaluation unit staff member at the time, and Alice Warner, who was a consultant to the Foundation helping organize the Devolution Initiative. It was the first step – described as a “learning manifesto” by Warner -- in developing a logic model for the Initiative, which would evolve as the initiative and evaluation evolved. The first of several versions of the logic model was developed several months later and included in an evaluation proposal by the Harvard Family Research Project (HFRP) to the Foundation in October 1997. Based on the statements of the premise paper and conversations with Foundation staff, HFRP set out the strategy and expected outcomes of the Devolution Initiative this way:

WKKF Devolution Initiative Learning and Logic Model



⁹ The first five grants for the Devolution Initiative were made in July 1996, before the “Premises Paper” was prepared.

The second source of early decisions about the Devolution Initiative and its evaluation was the deliberations and explorations of a cross-foundation team formed in 1996. Alice Warner described the team's starting point in this way: "We didn't do policy, we didn't do cross-foundation work, and we didn't do research." Nevertheless, a task force that included 15 to 20 program directors began a series of meetings to "see how to do cross-foundation programming" around devolution, according to Warner, and the group made decisions resulting in a July 1996 request to the Foundation's Board for the first of what would be four appropriations for the Devolution Initiative. Five grants were proposed, to support two nonprofit research institutions and three national advocacy organizations. The task force formed a "Core Team" for decisionmaking in late 1996.

In this early process of creating the Devolution Initiative, the decision was also made to manage it internally rather than to engage an intermediary organization – in part, to use the DI experience to practice cross-foundation work. Bill Richardson also had in mind that an internally managed initiative would maximize the Foundation's learning about devolution and maximize the Foundation's impact on the unfolding of devolution. To aid the internal thinking and management, consultants were brought in to advise on several aspects of the Initiative. These included specialists in policy and communications, who stayed with the DI over its course. Several evaluators were also brought in to participate in "framing" the initiative, including Heather Weiss of the Harvard Family Research Project, whose organization was selected in late 1997 to conduct an outside evaluation of the Initiative.

The Evaluation: A Focus on Learning about Grantees Informing Policy

The Harvard Family Research Project's evaluation design prioritized two goals – informing evolving Devolution Initiative strategy and fostering ongoing learning within the Foundation and with grantees. The primary audiences for evaluation findings were internal to the DI, rather than external audiences such as policymakers or the public. The DI evaluation approach emphasized examining results against the stated DI strategy for real-time learning. HFRP staff recognized that the devolution process would unfold without a predictable script and that new variables would have to be taken into account regularly. Rather than viewing this emergent process as a threat to evaluation – as it might have been in a conventional evaluation design of the time – the idea was to recognize the reality of change, and use the evaluation to stay on top of change and learn from it.

In its October 1997 proposal, the Harvard Family Research Project evaluators put forth the following questions for addressing the impact of the Devolution Initiative:

- (1) Is information reaching the intended audiences? Are different audiences finding it useful? What else do they need?
- (2) Are information providers adapting the type, form or timing of information in light of evolving customer needs? Is a critical mass of information available

- to meet customer needs and stimulate diverse stakeholder interest and engagement?
- (3) Is the information getting to intermediary organizations? Are intermediary organizations building the capacity of groups to use information?
 - (4) Are a range of groups getting and using information to inform policy discussion, development and modification?
 - (5) Is there evidence the Initiative is working so that policymakers are using information and stakeholder input to inform the policy agenda and decisions?

HFRP's evaluation used mixed (quantitative and qualitative) and multiple methods to assess the Initiative's complex layers. The evaluators also used triangulation – the use of multiple methods and/or data sources to study the same phenomenon – to examine the DI's intended outcomes and corroborate findings. A full list of evaluation methods and data sources is included in Appendix C.

The process of devising this evaluation plan and the process of carrying it out involved complicated issues of attribution. A question that kept coming up was: What should be learned about the “value added” of the Foundation's investments in the Devolution Initiative? At an early design stage, this question was discussed in terms of what was possible to learn about changes that could be attributed to DI. Then, in various discussions about how the evaluators should spend their time and effort, the question became: What should the evaluators try to learn about attribution? And at the end of the Initiative, the question asked – particularly by the Foundation's Board -- was: What was actually learned about the DI's contribution to policy?

Ricardo Millett, the Foundation's Director of Evaluation at that time, strongly believed that a conventional evaluation approach would not be useful in learning about the Devolution Initiative. The Foundation needed to understand “contribution, as opposed to attribution” – meaning that trying to link Foundation-funded activity to effects in the devolution policy landscape was important, but not as important as learning about “information sharing and points of connection,” according to Millett. Heather Weiss added credibility to the approach because of her strong reputation, Millett says.

The Harvard Family Research Project (HFRP) evaluators also felt that a fresh evaluation approach underscoring collaboration, continuous feedback, learning and flexibility was more appropriate than traditional evaluation approaches. As the HFRP proposal stated:

The nature of the Initiative – particularly its evolving character, diffuse “treatment” and probable “contamination” of the treatment by other non-WKKF information and capacity-building efforts – is such that experimental or comparison group designs are not feasible. Therefore, it will not be possible to make definitive causal statements about the impact of the Initiative. The proposed design provides data to test whether a *plausible* and *credible* case

can be made that the Initiative worked as modeled and positively affected the nature, availability and use of information as intended.¹⁰

The evaluators noted that attribution of changes in policy to the Devolution Initiative investments would not have been possible because several other national foundations were funding the DI grantee organizations to do the same or similar work across the country.

Bill Richardson was also disinclined to worry about attribution early on. He said, “If we had gotten bogged down at the beginning with the question of how do we know we’re making a difference, the Initiative would never have gotten off the ground.” Richardson was more concerned that the Foundation “keep track of what we’re doing” and learn whether some of the key policy-informing goals of the DI were being met, such as whether key national grantees were becoming capable of “working outside the Beltway.”

An evaluation design was chosen with both formative and summative components. In its formative approach, the evaluation was intended to provide timely, continuous feedback to the Foundation based on both qualitative and quantitative data collection about the initiative as a whole, including how grantee activities were -- or were not -- adding up to the DI strategy envisioned by the Foundation. In addition, the evaluation was intended to facilitate clarification of grantees’ goals; create a learning system among grantees, consultant partners, and the involved Foundation staff; and provide suggestions and assistance for “continuous improvement.” The evaluation was planned as a mechanism for strategic thinking about the Initiative throughout its course, to help the Foundation manage the initiative by feeding information to a Core Team making decisions about next steps. These evaluation goals were described in the October 1997 proposal to Kellogg for the evaluation from the Harvard Family Research Project team:

- (1) To work with WKKF and grantees to identify, refine, and modify important areas of inquiry and measurement that will inform the activities of the Initiative;
- (2) To provide baseline and continuous information about these measurement areas; and
- (3) To provide a mechanism for the timely flow of strategic information and feedback to support the development of the Initiative.¹¹

The Harvard Family Research Project team leaders characterized the situation going into the evaluation in this way:

¹⁰ Harvard Family Research Project. W. K. Kellogg Foundation Devolution Initiative Evaluation Proposal, October 10, 1997, p. 4.

¹¹ Harvard Family Research Project, p. 7. The proposal’s introduction to these three “interrelated” goals cites the December 18, 1996, paper by Kellogg Foundation staff known as the “premises paper.”

From the outset, the Initiative raised formidable challenges for both the Foundation and its evaluation. First, the cross-foundation team approach to grantmaking was completely uncharted territory. A process for collective and collaborative decisionmaking did not exist. Second, existing grants were retrofitted into a new initiative in which grantees were expected to work together, in addition to completing the terms of their original grants. The Foundation was challenged to make the whole of the Devolution Initiative greater than the sum of its individual parts. Evaluators were expected to help create as well as track “the whole.” Third, the funding of policy work and advocacy was a new direction for the Foundation. Fourth, there were many questions about how best to evaluate a complex initiative and particularly one that involved advocacy and policy. New approaches, such as theory of change evaluation, were emerging but had not been widely tested. Fifth, the approach of involving evaluators from the start was a relatively new one within the Foundation, particularly with respect to the evaluation’s relationship to the cross-foundation leadership team. And finally, all of this was taking place within an environment of restructuring and downsizing within the Foundation,¹² which meant that membership on the Foundation’s core leadership team shifted over the course of the Initiative.

One method of strategy support by the evaluators for the Initiative focused on reviewing and refining the DI logic model. According to the evaluators, this helped structure conversations in the DI core team about different ways in which the Initiative could develop. The logic model was revised several times, and elaborated, over the course of the Initiative. The underlying evolution of the Foundation team’s thinking confirmed for the evaluators the emerging nature of the Initiative as the devolution policy landscape clarified, and thus confirmed as well the need for a formative evaluation design. According to Heather Weiss, “We recognized at the outset that much about the Initiative would change over time, and that the evaluation would need to help inform those changes as well as adjust to them. We didn’t anticipate, however, the extent of the changes that would come.”

Some DI core team members felt too much time was spent on the logic model, however. Teresa Behrens, the current Director of Evaluation at the Foundation (who was not involved in the Initiative) saw the multiple revisions as evidence of lack of clarity about goals and methods. “It is not clear to me whether the revisions were based on new data or new thinking, or whether the multiple versions of the logic model were attempts to achieve consensus among program staff.” Anne Petersen, Senior Vice President for Programs¹³, agrees. She notes that a theory of change “is always flexible as things are learned in the course of doing an initiative, but not gaining agreement at the outset let it [the DI] be lots of things to different people. This was the Achilles heel for quite some time in the Initiative.”

¹² In 1998, the Foundation reduced its staff and grant portfolios by about a third due to declines in Kellogg company stock, which constituted about 70 percent of the Foundation’s assets.

¹³ Anne Petersen retired from the W. K. Kellogg Foundation in August, 2005.

Ricardo Millett admits that this observation is partly true, citing the advances in logic modeling for evaluation since 1997 as a positive development, as well as the Foundation staff's need in 1996 to support studies by several national research institutions of the effects of devolution. He also judges the evaluators to have been "clever" not to try to nail down a detailed theory of change at the beginning:

The original intent was to start with a theory of change design and have it be the driver of the evaluation. Heather [Weiss of HFRP] quickly realized that this conventional approach might not be seen as practical. So, instead of using a theory of change to understand and direct the implementation process, it became more of an explanatory paradigm that was fleshed out over time. Heather understood the actors and institutional context and was clever to shift.

In noting that "attribution" was not a primary objective of the evaluation, Millett also points to the evaluators' challenge to satisfy multiple Foundation clients: Even though Millett did not think an experimental evaluation approach made any sense for an initiative finding its way in an emerging policy landscape, the Foundation's co-chairs of the Devolution Initiative were frequently asked by Board members what difference DI was making and how the Foundation staff would know. According to Petersen, leaving the theory of change *too* flexible meant never being able to get to the intended impacts.

The summative component of the evaluation became more prominent late in the DI course in order to respond to this concern, but, in the view of several Foundation observers, it did not fully succeed in answering the question about the Initiative's impact. While some observers conclude that the evaluation emphasized learning about information flow and the utility of information for advocacy, rather than the results of information on policy, the evaluators disagree. And Millett concurs with the evaluators that they "tried hard to make the connection between the DI activities and policy outcomes." The Harvard Family Research Project evaluators cite their tracking of 22 policies in 1999-2000 and 47 policies in 2000-2001, combined with quantitative and qualitative state-level data collection, as a significant effort to determine if a plausible case could be made that the DI grantee efforts contributed to policy outcomes. The evaluators reported that the "intended policy results" were achieved for 17 of the 22 issues that grantees informed in 1999-2000 and 32 of 47 issues in 2000-2001.¹⁴ However, they were never invited by the Foundation to sum up the whole of the DI experience in an analysis that might be recognized as a typical final evaluation report.

In the late stages of the Devolution Initiative, the Foundation employed a communications strategy to answer the question of impact. For example, materials accompanying the final appropriation request to the Board for the Devolution Initiative told "stories" of new voices in state and local policy decisions and the human face of the Initiative – stories that were prepared by the Foundation's Communications unit, not the evaluators, although they were based upon evaluation learning reports and findings. The story of the Devolution Initiative for public audiences employed a communications strategy as well, which is discussed further below.

¹⁴ Harvard Family Research Project, *W. K. Kellogg Foundation's Devolution Initiative: Learning Report, 1997-2001 Evaluation Findings*, February 2002, p. 31.

Moving Parts and Moving Targets

Much of the evolution of the Initiative and its evaluation can be understood in the context of the Foundation staff's response to the shortcomings they came to recognize in the initial strategy – shortcomings that, in the view of the evaluators, *they* uncovered and brought to the attention of the DI core team. It became apparent that the funded national research institutions were accustomed to speaking to a national policy audience and were, in some cases, unsuited for the task of producing policy information useful for advocacy organizations, in general, and specifically for the state and local players in devolution. It also became apparent that national advocacy organizations, in some cases, did not have the resources, or the requisite relationships with each other or with their own local affiliates, to lead capacity building for effective state and local coalition building and citizen mobilization. In response, two years into the Initiative, the Foundation added a new set of grants and activities to build capacity for informing policy at the state and local levels of devolution. Further into the Initiative, a Scholar/Practitioner Program was added to support state and local policy research and serve as a policy training ground for minority scholars. (A description of the stages of the Initiative, including the shift to a state focus, begins on page 17 below.)

The result of these and other decisions was that the Initiative got complicated and, consequently, the evaluation of the Initiative got complicated. As several people interviewed for this case said, referring to either the Initiative or its evaluation -- or the two together: "It had a lot of moving parts." In the end, the complicated DI strategy and an evaluation designed to keep track of all the moving parts – but not to sum up its impact -- contributed to an indistinct "bottom line." Petersen notes that the evaluation should have helped to clarify, but instead added to the complexity.

A separate internal evaluation of the Devolution Initiative provides information about how this experience affected the Foundation. As the Foundation's first significant cross-program initiative, the first significant policy-focused initiative, and the first major initiative evaluation not managed within a program division, the Devolution Initiative introduced some ways of doing business that have had lasting effects. Lessons documented in two internal evaluation reports are drawn from for this case, along with staff observations, to illuminate how the DI served the purposes of the Foundation's leadership to restructure and strengthen the philanthropy's ways of working.

Background: The Foundation People and the Institution

At the time the Devolution Initiative was being formed, the Foundation had four major program areas: Health, Youth and Education, Philanthropy and Volunteerism, and Food Systems and Rural Development. Foundation President Richardson had also instituted a "matrix" structure with themes that cross-cut these four program areas, including Leadership, Social and Economic Community Development, Capitalizing on Diversity, and Information and Communication Technology. In addition, there were geographically defined programs for Greater Battle Creek (MI), Southern Africa, and Latin America and the Caribbean.

A key organizational decision for the Initiative was the naming of C. Patrick Babcock as the Foundation's Director of Policy. Babcock had joined the Kellogg Foundation in 1991 as a program director in Health programs after serving as the director of the Michigan Department of Social Services. In 1995 he became the director of the Foundation's new Policy department, one of four parts of a new "Impact Services unit," which also included the Foundation's Marketing and Communications, Technology, and Evaluation departments. The unit's name was intended to signal a commitment to increasing the impact of program activities. (The unit has since been expanded to include two additional departments, Organizational Services and Program Learning.)

Babcock became co-chair of the Devolution Initiative with a long-time Foundation vice president, Dan Moore, and brought to the Initiative the perspective of what is needed to affect state and local public policy. Moore was Vice President for Philanthropy and Voluntarism at the time, having held three different Vice Presidential jobs. Both co-chairs were experienced bridge builders, according to Moore, and both were committed to a participatory, consensus-based approach to running the Devolution Initiative.

The Impact Services unit reported to Senior Vice President Anne Petersen, as did all regular program areas. Petersen was brought into the Foundation by Richardson in 1996, shortly after the Devolution Initiative got started. The Devolution Initiative, however, was a cross-program operation, staffed by program directors (Kellogg's equivalent of "program officers") from three of the four program areas. Initially, budgets for DI grantmaking were under the authority of the cross-foundation team, while budgets supporting program director staff remained under the control of program areas.¹⁵ In recognition of the administrative challenges involved in a cross-program initiative that was expected to be significant, a consultant who assisted Pat Babcock, Alice Warner, was hired to provide full-time support to the Initiative; she was joined in 1999 by a program assistant, Diane Smith. After Smith joined the DI group, she became solely responsible for tracking all DI grants, contracts, and budgets centrally, regardless of where the grants were "housed" or which program director was responsible.

Ricardo Millett headed the Evaluation department that, after Richardson took over the Foundation, became part of Impact Services, and he continued in that role until April 2001. Millett reports that Kellogg was an early leader in philanthropic evaluation with the "insight to give credence to program evaluation." But, he says, implementing the commitment was difficult. The historical problems included typical foundation struggles for control of resources. The evolution of the evaluation function at Kellogg was not unlike what was going on in foundation evaluation units across philanthropy, according to Dan Moore, who served as the Foundation's Vice President for Strategic Planning, Evaluation and Technology in 1990 and 1991. Ironically, while it was difficult to implement evaluation inside the Foundation, the Foundation's reputation as a leader in

¹⁵ As described in the internal evaluation of the Devolution Initiative, the financial and grantmaking "architecture" of the Foundation was not suited for cross-program activity and complicated the work of the DI team considerably.

program evaluation outside the Foundation was elevated by two popular evaluation manuals published on the Foundation's website during Millett's tenure.^{16 17}

Inside the Foundation, there was concern on the part of program staff that the evaluation unit aspired to independence. Working against this was a structure in which program areas controlled all evaluation funding except for staff-related costs of the evaluation unit. They also selected and managed evaluators of projects and did not share the results of project evaluations with the evaluation unit. In Millett's description, evaluation was a foundation-wide service to programs that was "rejectable" – meaning that program staff could accept or reject it as they chose. On the other hand, program vice presidents acted to create new evaluation manager positions during the process of downsizing that occurred in 1998, according to Anne Petersen.

The current Evaluation Director at the Foundation, Teresa Behrens, says that this structure was not the only reason that program areas were deeply involved in evaluations. The evaluation unit was never fully staffed until late 2003, according to Behrens, so "program directors had to be more involved in controlling the content of evaluation because there wasn't the expertise in the evaluation department of Impact Services. [That department] had a 'chicken and egg' problem of staff turnover: There were not enough staff, so the existing staff were too stretched and were not able to do the job well."

Anne Petersen saw resistance to the accountability of evaluation. She observes that the Foundation's evaluation work has historically focused on process: "The motto was 'improve, not prove,' but there was no structure for learning from even the process evaluations – the 'improve' part." Further, Petersen observes that "there was never a serious interest in results-oriented evaluation at the Foundation... There's a fear of assessment or fear of feedback. It's a lack of trust in how it's going to be used." With respect to the DI evaluation, Petersen reported that "the view of a lot of staff was to keep it as far away from them as possible."

Millett saw progress toward resolving the typical program *versus* evaluation tensions in the years just before the Devolution Initiative and this progress has since continued, according to Behrens. Petersen concurs, crediting both Behrens and a greater

¹⁶ W. K. Kellogg Foundation, *Logic Model Development Guide: Using Logic Models to Bring Together Planning, Evaluation & Action*, updated December 2001, and W. K. Kellogg Foundation, *Evaluation Handbook*, January 1998.

¹⁷ Another of the legacies of evaluation at the W. K. Kellogg Foundation was the concept of "cluster evaluation." The term was adopted in 1988 by a previous Kellogg evaluation director to describe the practice of evaluating collections of projects with similar goals. This practice shaped the development of the Devolution Initiative, although there are different views of what "cluster evaluation" is. In one view, "The Kellogg Foundation continues to use this [cluster evaluation] approach, now emphasizing systemic and policy-change outcomes, as well as the process for identifying common themes in retrospect, after grants have been made." Ross F. Conner, Victor Kuo, Marli S. Melton, and Ricardo A. Millett. "Adapting Evaluation to Accommodate Foundations' Structural and Cultural Characteristics" in Marc T. Braverman, Norman A. Constantine, Jana Kay Slater (eds.), *Foundations and Evaluation: Contexts and Practices for Effective Philanthropy*, San Francisco: Jossey-Bass, 2004, p. 172. Teresa Behrens, the current Director of Evaluation at the Kellogg Foundation, thinks that the DI – and therefore, its evaluation -- was more like a cluster of similar projects and less like a unified initiative than initially conceived.

insistence on evaluation from the Board. The evaluation department developed an internal contracting procedure that outlined the roles and responsibilities of the unit, which included the selection of evaluators, and co-management of evaluation budgets with program directors, for example. The Foundation expectation, going into the DI, was that major initiatives for Board review and approval would be evaluated and that initiative leaders had to articulate expected outcomes in the initial Board materials.

The Devolution Initiative was an opportunity for the Evaluation department to put into practice some of the beliefs about program effectiveness espoused in the Kellogg-published manuals and in technical assistance provided to other foundations using these materials. It was also an unusual opportunity to manage an evaluation. According to Alice Warner, Millett “held the contract” for the DI evaluation, which would not have happened if the DI had been housed within a program area. (Millett acted as the liaison between the Harvard Family Research Project evaluation team and the Foundation until his departure in April 2001, when the evaluation team began reporting to Pat Babcock and Alice Warner in the Foundation’s Policy department.) For the Devolution Initiative, the evaluation unit “preached early [evaluation] involvement in the initiative development process and theories of change to help frame evaluation designs,” according to Millett. He said this approach was generally accepted, at least at the beginning of the DI process, because it was championed by Pat Babcock, who was highly regarded in the Foundation. In some ways, however, the resulting evaluation was a compromise from Millett’s perspective, as discussed below. (See “Views of the Devolution Initiative Evaluation.”)

Also important to the Devolution Initiative was the Foundation’s Marketing and Communications department within the Impact Services unit. Headed by Karen Lake, the department managed a national opinion survey and the “roll out” of survey findings that launched the second – state and local – phase of the Devolution Initiative; designed and operated a comprehensive website that featured hundreds of articles and materials published on devolution-related topics by participating grantees; and published a book to tell the DI story to external audiences, entitled *Building Bridges Between Policy and People: Devolution in Practice, Lessons from the W. K. Kellogg Foundation’s Devolution Initiative*. The Marketing and Communications department also assembled a “disseminators group” among the DI grantees to collectively advance the goals of the Initiative through information strategies and to “avoid redundancy and unnecessary use of funds,” according to Lake. Importantly, Lake says, the group included members of the policy and evaluation teams, including all consultant groups that supported these areas. In Lake’s view, their involvement helped to keep efforts closely aligned and in support of one another.

Not unlike the evolution of evaluation at the Foundation, the communications role *vis a vis* program shifted over time. Historically, the unit housing communications looked at dissemination only, and did this at the end point of projects for project reports, and annual reports. The shift was to “having Communications at the table at the beginning of an initiative,” Lake says, beginning with a Community Partners in Health Professions initiative in the early 1990s. Lake says the idea in this shift was to ask, “What is it we’re trying to change and how can we use these tools – communications, evaluation, policy, and technology – to leverage impact?” Communications played this

at-the-table-early role in the Devolution Initiative, participating in the Core Team that managed the Initiative and served as the executive decisionmaking body. How much of the Foundation’s final story about the Devolution Initiative was shaped by an evaluation perspective and how much by a communications perspective – where both were involved – is a key issue featured in this case.

The Evaluators

The Devolution Initiative Team selected the Harvard Family Research Project (HFRP) as the evaluator for the initiative, with Ricardo Millett serving as the primary advocate for HFRP. While other evaluation groups were invited to propose their ideas for the DI evaluation, HFRP, led by Heather Weiss, persuaded the DI Core Team that they understood what was being asked for and that they could deliver it. In part, this occurred over the course of a year-long consultancy by Weiss to the Team before HFRP was selected as the DI’s outside evaluator. According to Weiss, Millett wanted a group that knew the substantive issues as well as evaluation, an approach that supported real-time learning, and people who could work within the “shared governance” structure of the DI and its evaluation. (From Millett’s perspective, it was clear that the challenges would require evaluators with “a lot of patience.”) Pat Babcock wanted an evaluator with both quantitative and qualitative skills, a track record of innovation, and the ability to manage “lots of moving parts.” “We were looking for someone who could give us their best judgment,” Babcock said.

In its October 1997 proposal, the HFRP offered a multi-dimensional set of services to the Foundation:

- Tracking and documentation of information flows to learn how and with what content the grantees were reaching their target audiences;
- An analytic component to provide information about whether the documented information flows were likely to achieve the Foundation goals of informed policymakers and informed and empowered citizens;
- A theoretical component, offering a framework for effective application of knowledge to the policy process;
- A strategic development component, embedded in a process of unpacking and helping Foundation staff articulate and refine their assumptions about grantee activities and effects -- as expressed in an evolving logic model;
- A learning system building component, facilitating meetings of grantees and Foundation staff to review events, understandings, and lessons – particularly by providing lesson-oriented documents in advance of these meetings; and
- A measurement component, designed both to “count” the output of the DI grantees in meaningful ways and to assess the effects of grantee activities on policy based on a standard of plausibility (rather than proven attribution).

HFRP proposed to stage these activities over the five-year course of the Initiative, emphasizing relationships for strategy and network building early, measurement late, and tracking and documentation throughout. Ultimately, the evaluators produced more than 40 documents for the Foundation about the Devolution Initiative, sequenced with data collection and reporting to inform Initiative decisionmaking and learning needs. (See Appendix A for a list of evaluation reports and their timing.)

Weiss and her team's experience in formative evaluation not only involved a high degree of communication with program people unusual in evaluation practice, but an active role in helping program people formulate their theories of change. This role would come to look like a seat at the decisionmaking table to some Foundation participants. But to others, it looked like the evaluators "putting the Foundation staffs' feet to the fire" about what they wanted to accomplish and, to still others, it looked like the evaluators bearing bad news or opening up topics on which a consensus had not been reached among the DI Core Team members.

The evaluation team eventually included Weiss as the Principal Investigator, Julia Coffman as the evaluation's Project Manager, five senior Research Associates and two Research Assistants. Team members worked close to full time on the Devolution Initiative evaluation.

The Major Stages of the Devolution Initiative

After the creation of a Task Force and then a smaller Core Team, and recommendations for the first five grants to the Foundation Board, all in 1996, the Initiative shifted to a stage of trying to understand how to connect the national grantees funded in the first and second round of DI grants to state- and community-level organizations that were, theoretically, at the front of devolution. According to Heather Weiss and Julia Coffman of HFRP, "there was lot of discussion about how to 'drill down' from the national level, or how to make sure that the Initiative would 'have legs.' It was in the logic model, but not in their grantmaking." At the September 1998 annual networking meeting of grantees, Foundation staff and consultants, which was held at Gallaudet University in Washington, D.C., discussion focused on "where does DI hit the ground at the state and community level?"

According to the evaluators, when Dan Moore, Pat Babcock and Ricardo Millett brought the Gallaudet discussion back to the Foundation, there was a "go" for adding a new set of activities to produce, disseminate, and help users work with information for policy advocacy in a concentrated effort in selected states. Organizations in Florida, Mississippi, Montana, New York, Washington, and Wisconsin were selected for state-level grants to ensure that the Devolution Initiative's investment in producing information would, indeed, "hit the ground."¹⁸ Foundation program directors were assigned to oversee the state-level investments and the Harvard Family Research Project added evaluation methodology to track state- and local-level investments and outcomes. HFRP field staff were assigned to track the activities of the state- and local-level grantees

¹⁸ Florida, Mississippi, New York, Washington, and Wisconsin were known as "focus states." In Montana, a nonprofit advocacy organization worked with the state's seven tribal governments.

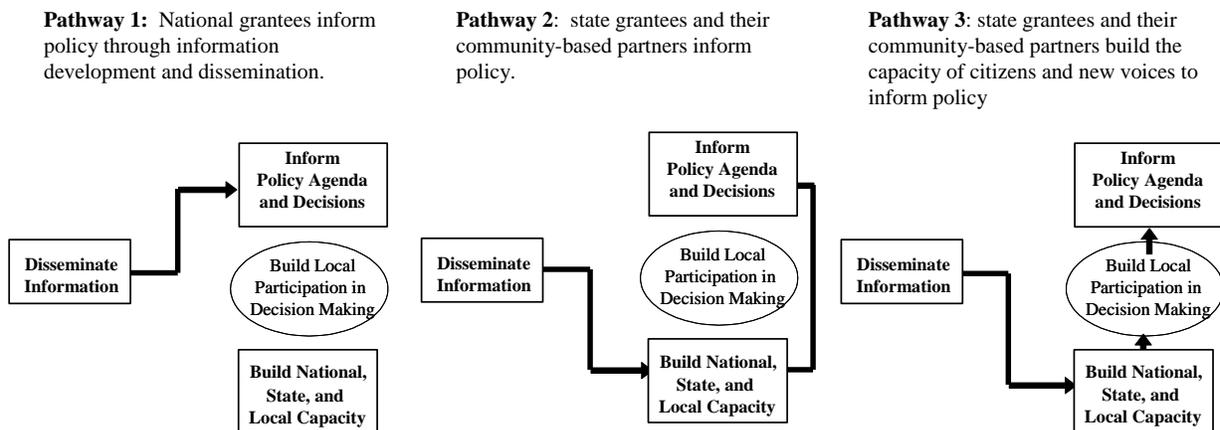
and, therefore, in addition to tracking and reporting on the work of the overall Initiative, evaluators began reporting separately on work and outcomes in each of the states selected for focused attention.

Then, according to the evaluators, a year-long discussion ensued among the Foundation players about community mobilization theories. There were multiple points of view across the Devolution Initiative Core Team. This discussion showed up in changes to the DI logic model with the addition of a new objective -- to “Include New Voices” in the policy process.

In 2000, evaluation results from the national partners began to highlight health disparities, and economic and social disparities, across population groups in the U.S. Data from the Immigrant Welfare Implementation Collaborative, the Urban Institute, and the Joint Center on Political and Economic Studies, for example, spurred those grantees and others to start looking at the effects of devolution on Native Americans, immigrants and other groups, and provided the content and direction for the work of the Scholar/Practitioner groups in two states (New York and Mississippi).

Another shift occurred when the Core Team realized that a push to generate and disseminate information for state-level policymaking would not necessarily generate “new voices” in policy discussions and new levels of engagement. As the evaluators tracked the implementation of the DI theory of change implicit in the logic model, it became clear that DI investments represented three major “pathways” through the logic model to informing policy.

Pathways to Informing Policy



Pathway 1 was the most direct route to informing policy, with a line from grantees (in this case, national grantees) and their information to informing policy. Pathway 2 also informed policy, but reflected an investment in building unified voices among state and community advocates to do it. Later DI investments in state grantees and their community-based partners “gave legs” to and pushed for results along Pathway 2. Pathway 3 “drilled down” deeper into communities and reflected an investment in state grantees and their community-based partners to build capacity of community members, particularly new voices underrepresented in the policy process, to inform policy. (These pathways were unpacked from a comprehensive model that appears in Appendix D.)

Foundation program directors disagreed on which pathway should be a priority for the Devolution Initiative. Specifically, there was a lack of consensus about whether it should be a priority to encourage and empower new participants in policy discussions at the community and state level (Pathway 3), or whether the priority was to actually influence state-level policy (Pathway 2), which efforts to include new participants might dilute. The lack of consensus on this point led to the Foundation’s state-assigned program directors each “doing their own thing,” according to the evaluators.

As the national part of Initiative was drawing to a close in 2001, Pat Babcock and Alice Warner asked the HFRP team to focus on the sustainability of the processes and organizations initiated under DI to continue to inform policy, and measurement of the Initiative’s effects. (State-level DI work continued for an additional two years with supplemental funding from the Foundation.) The evaluators fielded a survey at this stage to document the extent to which state-level policymakers attended to DI grantees’ communications. They also developed “stories” using human interest journalism techniques to communicate how devolution was “hitting the ground” in the states selected for the Initiative. The evaluators summed up findings for the three pathways to informing policy in a series of reports in November 2001,¹⁹ which were consolidated and summarized for a final briefing of Foundation staff by the evaluators in February 2002. These reports were prepared for an internal audience – i.e., the Foundation staff and grantees.

In a separate effort to tell the DI story at this stage – preparing for an end-of-project conference with the national grantees originally scheduled for October 2001 (and rescheduled after the events of September 11 to be held in January 2002) -- the Foundation’s Marketing and Communications department set out to produce a book to explain the Devolution Initiative to the policy and public audiences, which was published December 1, 2001.

Views of the Devolution Initiative Evaluation

In general, the Foundation staff interviewed for this case found the DI evaluation helpful and enriching to the implementation of the Initiative. But it was also complicated and organizationally difficult. Miguel Satut, a program director from the Foundation’s Youth and Education area who was responsible for overseeing DI implementation in

¹⁹ Harvard Family Research Project. *Pathway 1: 1997-2001 Findings and Implications; Pathway 2: 1999-2001 Findings and Implications; Pathway 3: 1999-2001 Findings and Implications*, November 2001.

Wisconsin, said “the added value of the Harvard Family Research Project was the larger picture of the whole initiative.” Barbara Sabol, a program director in Health, appreciated having “a different pair of eyes” on what was happening in devolution and among the DI grantees in Mississippi, the state that was her responsibility. “Knowing that there was an evaluation going on reinforced [Sabol’s intent] to address issues, probe, explore, provide assistance.” The framework of DI and its evaluation provided an opportunity for Mississippi to “get highlighted for something other than being at the bottom of the pile.” Sabol’s added-value list for the evaluation includes:

- Evidence from a neutral third party;
- Questions that precipitated thinking and re-thinking;
- Selective reading of documents for information important to individual states; and
- Contribution to cohesiveness of the DI group.

On the other hand, according to Sabol, it was expensive – “there was a question of balance.”

Allocation of resources within the Devolution Initiative was a common criticism from program directors who had state responsibilities. Their view was that national grantees received large grants early in the Initiative, but when it was determined that information production, dissemination and coalition building were needed at the state, local, and tribal levels of government, the resources did not follow. In contrast, the evaluators reported that resources were adequate for them to follow the DI action as the number of grantees expanded, the geography changed, and the focus of their activity shifted from formative feedback to outcome measurement.

Henrie Treadwell, a Health area program director responsible for overseeing DI in Florida, supported the formative approach in terms of the feedback provided – “there was no other way to do it” – and believed that the evaluation “made the process more rich. We had to be thoughtful, analytical throughout, willing to reflect and act.” She was not as appreciative of the evaluation design, however. To Treadwell, “the evaluation wasn’t conceptual. It was not based on ‘Here’s what we need to know’ but on ‘Go look around.’” In answer to her own question of “Would it have been possible to specify [the evaluation] more carefully?” Treadwell said, “The table wasn’t set for that discussion on either side – to push on what do you really want to know. The process was arduous.”

The mechanism for team learning most often mentioned by Foundation people interviewed for this case was a quarterly “learning meeting” of the Foundation’s Core Team facilitated by the evaluators. To support the evaluation’s formative component, the Harvard Family Research Project designed these meetings as an opportunity to convene the Initiative’s Core leadership team for “real-time” discussion about what was being learned from evaluation data collection on topics that had particular relevance to upcoming strategic decisions. Meeting topics were chosen in conjunction with the Foundation at the start of each year and the HFRP evaluators distributed briefs averaging about 15 pages at least one week in advance of the meetings. (Appendix B provides a list of materials prepared for these meetings.)

While the learning meetings were pointed to as an innovation, the discussions were not as effective as they might have been, according to Millett, Warner, and Babcock. According to Warner, “It took a long time to do it well.” Pat Babcock’s view is that in the early stages of state-level work, the Foundation’s program directors knew more, sooner about what was happening on the ground than the evaluators were able to bring to the discussions; the written materials prepared by the evaluators for the meetings were often distributed at the last minute; and the “team style of meeting conflicted with the fact that content was being brought by the evaluators” – non-team members playing an expert role. “We needed better ground rules going into the discussions,” he reflected.

In the process of getting these meetings to work well, some Core Team members formed the view that the evaluators were inappropriately involved in decisionmaking. According to Millett:

The structure of the external evaluation drove the management of the initiative because staff came from all over the Foundation and it was so hard to [achieve consensus]. Lack of time by the Core Team was an issue.

Heather Weiss and Julia Coffman disagree with this assessment. In their view, the process of pushing Foundation staff to articulate and examine their assumptions about how DI activities would affect policy might have given some Core Team members the impression that the evaluators were in the lead, but the evaluators were only doing their clarifying job. From the evaluators’ perspective, they advised the Foundation based on evaluation data collected, but the Core team made all Initiative-related decisions. Pat Babcock concurs. In his view, it was the role of the evaluators to ask hard questions. The result was sometimes that this process kept resurfacing issues on which team members had “agreed to disagree,” which some participants did not welcome. Babcock’s priorities about the credibility of “arms length” evaluation *versus* involvement were stated this way: “I would rather take the risk of loss of credibility than have an evaluator outside the loop.” According to Dan Moore, “the DI unfolding was the driver... The strategies of the evaluators to help the Core Team draw out their assumptions and logic models worked to force clarity.”

Ricardo Millett observed that the DI evaluation sometimes was a compromise. In order to get buy-in for the evaluation in the first place, it had to be more process-oriented than outcome-oriented, which was eventually a drawback in terms of credibility, although “ultimately, we made the right decision [about choosing the Harvard Family Research Project evaluation team]. It was a tough piece of work.”

Teri Behrens adds to these views an observation of the *détente* phenomenon: In the absence of an overarching structure for resolving differences, senior foundation staff not only agree to disagree, but they work out the boundaries of their autonomy so that the practice of grantmaking within a single organization may differ from area to area by informal agreement.

In spite of a favorable overall assessment of the Devolution Initiative evaluation by most of the involved Foundation staff, Millett was disappointed, as were others, with

the Initiative and the evaluation in terms of the diversity of perspectives included. Professor Ronald Walters of the University of Maryland's James MacGregor Burns Academy of Leadership was brought in to advise the Core Team and the evaluators about reaching minority audiences with policy information, and this effort was judged to be effective in the early stages of polling and then in shifting the Initiative to a state- and local-focused strategy. Walters was subsequently engaged to head a Scholar/Practitioner Program, which was intended to draw minority scholars into DI research on policy topics at the state and local levels of the Initiative. Each state (except Montana) had a Scholar/Practitioner team led by a university-based faculty member (known as a "mentor"), whose activities were coordinated by Walters. Further, the Foundation pushed HFRP to diversify its research team. Neither of these efforts to diversify perspectives was deemed successful by the people interviewed for this case. (Walters was not interviewed.)

The relationship between Walters' Policy Think consulting group and the HFRP was a "point of tension," according to Millett and there were "HFRP compromises on getting an authentic view from minority communities." There was concern expressed by the Foundation's program director for the Florida DI activities, Henrie Treadwell, that it might not be possible for a white field researcher to ascertain how devolution was affecting citizens, particularly people of color, in Miami-Dade County. Treadwell wanted to be sure all voices were being heard in policy discussions. (Treadwell expressed similar concerns about the racial imbalance in the leadership of one of the Florida DI grantees.) Further adding to these tensions was the fact that the person hired by HFRP in response to Foundation pressure to diversify its research team worked on the evaluation for about two years but did not stay with the organization.

Finally, contributing to the diversity disappointments of the Devolution Initiative, the Scholar/Practitioner Program was judged an overall disappointment both because it did not prove to be a route for minority scholars to join the national research grantee organizations and because some of the SPP teams worked on independent projects of their own devising rather than projects connected to the DI activities going in the states. Barbara Sabol, the Foundation's program director for DI activities in Mississippi found the SPP in that state an asset because it produced state-specific research and included community voices in dissemination of the research findings. Pat Babcock notes that the Scholar/Practitioner Program contributions in Mississippi and New York State, and in Washington State at the end of the Initiative, were most closely related to the work of state grantees. Overall, however, he emphasizes the lack of diversity within the DI as a problem throughout its course, and the failure of the Foundation's attempts to address the problem.

The "Bottom Line" and the Devolution Initiative Legacy

The Devolution Initiative changed the way the Kellogg Foundation works in many ways. The program directors overseeing state DI work took lessons from the DI experience into new initiatives and projects. For example, Miguel Satut adapted DI findings in the Foundation's ENLACE initiative (Engaging Latino Communities for Education), another project "with lots of moving parts."

When you do state work, the focus has to be very clear to all partners. You need to convene everyone to think about the whole state and you need to recognize that there are different central, northern and southern cultures [in New Mexico] as well as different histories and conditions. You need to work intensely and with forethought about how to make it all cohere...Don't assume that for national organizations with state affiliates or chapters, they know at each level what is going on at the other level. Even in the very best national organizations there is a lag time for information exchange – and by the time the information gets there, state conditions may have changed.

Within the Foundation, policy work is accepted as a given now. According to Dan Moore, “policy comes flowing off people’s tongues more easily” and many Foundation initiatives involve policy work. In addition to ENLACE, the SPARK school readiness initiative has a big policy component and work with school health systems, work in philanthropy calling for accountability, and work in the Food Systems and Rural Development area involve policy activities and outcomes. Alice Warner named the Community Voices health care initiative for the underserved and the Turning Point public health collaboration with the Robert Wood Johnson Foundation as examples of Kellogg Foundation projects that were based on the state capacity built during the Devolution Initiative; states involved in DI were selected for these new initiatives because there was known capacity in them. Pat Babcock notes that the Devolution Initiative was one of the few times that the Foundation was “in the middle of the active policy development process”; usually Foundation projects looked at policy earlier or later.

Anne Petersen says that the Devolution Initiative helped the Kellogg Foundation find its policy niche and that Pat Babcock deserves the credit for helping the Foundation staff and Board see that “public will building” is that niche.

Before the Devolution Initiative, everyone thought of policy as something that happened at a very high level – something that policy wonks and policymakers did. From DI, “public will building” emerged as a more acceptable label for voice, participation and advocacy – the idea that people in a community should be able to say what they want to have happen. That fits with [the WKKF mission of] helping people help themselves.

Further, understanding and articulating the Foundation’s policy niche in this way has brought clarity to staff and Board assumptions about how information matters, according to Petersen:

The fuzziness in the concepts about what we were aiming for was the main problem with the Devolution Initiative evaluation. And, because some of these things are politically charged, there was a little of “working under the radar” – because the Board was uncertain about the value of policy research and did not want to fund research. Now, more people see information as a tool for change.

The legacy of the DI evaluation within the Foundation is a strong commitment to theories of change. According to Teri Behrens, however, one of the unresolved evaluation issues in the Foundation is: What *is* a theory of change – as opposed to a logic

model. According to Anne Petersen, the Kellogg Foundation Board likes and insists on the theories of change approach. In addition, the Board has approved a staged approach to designing initiatives that is intended to produce a strong, detailed theory of change at the end of an initial exploratory stage – before requests for major investments – partly in reaction to what Petersen and some Board members considered vague expectations of the Devolution Initiative and its evaluation. According to Pat Babcock, the Devolution Initiative had a “middle ground” theory of change as the strategy evolved for information development, dissemination and voice/participation. “The [starting] logic model was too generic; it was useful to explain what we were trying to do, but we added to it in discussion. It was not fully refined, articulated or even fully developed [at the beginning of the DI].”

Alice Warner reports that the Foundation is doing much more project and cluster-focused cross-foundation work as a result of the DI experience, but was just considering another initiative-level piece of cross-foundation work in 2004. She says, “The DI, in influencing many programming processes, has created a template for many initiative design features. Learning from the DI is helping us focus our knowledge management more aggressively, and figured into the systems design of finance/operations technology as well as the programming systems for technology.”

Pat Babcock credits the DI experience with several types of management learning, including lessons on how to manage a large, multi-stage initiative and the level of resources needed for such ambitious undertakings. If he were to do something like the DI again, he would break the evaluation up into several contracts to more closely fit with the different types of learning needed, for example. The DI confirmed for the Foundation, in Babcock’s view, the value of having an integrated Impact Services Team involved in major initiatives, as well as the need for identifying a core team and for practicing team decisionmaking.

Bill Richardson sees the Devolution Initiative experience as a bridge for the Foundation to new ways of working. It demonstrated how the Foundation can be a catalyst for community mobilization and awareness-raising; heightened the Board’s interest in both policy and evaluation; increased the Foundation’s evaluation expertise; and provided the model for the organization to work through cross-program initiatives.

The difference that the Devolution Initiative made in terms of social policy is not so clear, however. This is true for two main reasons. First, while the summative component of the external DI evaluation did measure outcomes, such as the number of grantees’ information products, and did attempt to make a credible and plausible case for how DI activity at the state and local level influenced policy, there were questions about its design and results. A survey of state decisionmakers got a poor response rate, according to Pat Babcock, because it was “not tied to something recognizable to the prospective respondents” (such as a state university policy center, instead of the Harvard Family Research Project for the Kellogg Foundation’s Devolution Initiative). As a result, the survey results were not credible, according to Babcock, and the evaluation was otherwise “relatively short on quantitative data.”

The evaluators point to their multi-method approach and concluding summary reports to counter this assessment, but several people interviewed for this case, including the evaluators, reported that the Initiative lost momentum in its final stage and the Foundation participants were not have been focused on the “summing up.” The Foundation participants were turning their attention to new projects, the events of September 11, 2001, caused the final meeting of national grantees to be delayed, and, while the evaluators had been focused on the Foundation and the DI grantees as their learning audience, the Foundation needed to tell the DI story to a public audience toward the end of the Initiative. In addition, from the evaluators’ perspective, the evaluation lost its internal advocate within the Foundation when Ricardo Millett left in April 2001. For all of these reasons, the Foundation did not ask the evaluators to produce the type of final report typical for projects of the magnitude of the DI – a document bringing together findings and lessons on implementation, impact findings, and reflection and interpretation of the contribution of the Initiative in the policy context of the time.

The Foundation made the decision to produce in book form its own version of the DI story for the public audience. According to Alice Warner,

The book was seen as a final summary of “what has been learned” in the Devolution Initiative. We found that we had done a lot of work (including the evaluation) on process, and yet the development and dissemination of a massive amount of information (I think between 2000-3000 reports, published documents, etc.) was produced or co-produced by various grantees with research on content. Two major pieces summarized the learning – the topic issues papers (an amazing collaborative writing process across organizations with dissimilar topic agendas) and the book.

The topic issues papers referred to by Warner were published as appendices to “the book,” *Building Bridges Between Policy and People: Devolution in Practice*. Warner further reported a large mailing of the book to policymakers, along with publicity by grantees; it is also available, chapter by chapter, on the W. K. Kellogg Foundation website and was promoted on grantee websites and in various listservs.

According to Karen Lake, the *Building Bridges* book was conceived “to raise the visibility of the DI work, specifically the way it went about to address the issues. It was not a replacement for evaluation, but was intended to draw upon some of the knowledge, experience, etc., gleaned through the process.” Further, Lake says, “we wanted to be ready for the closing conference [of national grantees] so that we had something meaningful to distribute that documented the long journey from the grantees’ perspective.”

Ricardo Millett and the evaluators judged the book to be a substitute for an evaluation-oriented summing up. On the other hand, it is not clear what an evaluation-oriented summing up would have been able to deliver, according to Anne Petersen, who said, “When evaluations aren’t designed to deliver the bottom line, foundations don’t have much choice.” About the DI evaluation she says, “The outcome evaluation had some results, but it was largely anecdotal – which makes it difficult to know how to improve.”

Alice Warner, who is widely known in the Foundation as the “institutional memory” of the Devolution Initiative, comments on the attempt to characterize a DI bottom line in this way:

So much of this retrospective suffers from the second hand quarterbacking – what could have/should have been done. The field of evaluation and many of the methods used for policy-related research measurements were in very different places within their own cultures when we started. More evaluators of true systems work (where everything is messy, interrelated, and everything moves) have moved thinking light years in the past few years. Why this is key is that the sustainable part [of the DI] is the relationships among and within the coalitions and alliances. Those are still working. Some institutions [among the DI grantees] changed internal policy [and some individuals from those institutions] have taken these lessons into new institutions—and I can already see the application of lessons learned. So, how does one measure relationship like that? What are the stages? And, where do you stop seeing the ripples?

Epilogue: The Internal Evaluations

Toward the end of the Devolution Initiative, the W. K. Kellogg Foundation instituted a Planning, Programming, Management system for funded projects and initiatives that describes best practices and requirements for each stage of grantmaking activity, including close-out – a system that Alice Warner says modeled its initiative processes on the Devolution Initiative. In this system, the Leveraging Impact and Learning (LIL) process is a requirement for closing initiatives that recognizes and formalizes the Foundation’s desire to be a learning organization. The Devolution Initiative was the first Foundation initiative to be studied in this LIL process to “harvest” its lessons for future work. The LIL study of the DI was the first of two internal evaluations.

The idea of a Leveraging Impact and Learning study is to “identify transferable concepts, strategies, and practices and to share insights with selected external audiences,” according to the authors of the “Final Report: Leveraging Impact and Learning from the Devolution Initiative of the W. K. Kellogg Foundation,” (March 25, 2004) by Shelley Stark and Mark Lelle, Foundation consultants. Drawing from multiple sources, including the Harvard Family Research Project evaluation reports on the Initiative, this document provides succinct, definitive statements of DI results, including the following (which have been edited for length):

- The DI partnerships between national and state grantees kept the pressure on the entire system for accountability to those in communities, which was helpful in maintaining supports when state and federal budget crises were forcing cuts in many programs.
- The initiative produced data on implementation issues needed to create a reinforcing system of policy for social change.

- Both the policymakers and the general public who interacted with them changed behaviors in ways that supported social change reflective of experiences in their states.
- State grantee lead organizations formed broad based coalitions and partnerships that included many people affected by the policy changes, but not at the table for policy decisions and all state grantees developed community leaders to carry on the policy change efforts for the future.
- Information development for a range of audiences provided multiple places of entry into the policy system as well as promoted grantees thinking through positions for messages for maximum impact for the general public.
- The requirement of work and inclusion of new voices in state efforts resulted in changes in governance structures of grantee organizations to include those being served as members of grantees' organizations.

The Devolution Initiative LIL study was the first and so far only such Foundation review of an initiative and, for Teri Behrens, it raises questions about who should do the LIL studies and how these should be managed. For example, one question is whether harvesting the lessons of an initiative should be the final responsibility of the program staff involved or whether this responsibility should be assigned to external people more removed from the experience. In the case of the Devolution Initiative, a middle ground was chosen – one consultant who knew both the Foundation and the Initiative's beginnings, having been a Foundation employee at that time, and one who had worked on the Harvard Family Research Project evaluation.

The Foundation judgment for the LIL study of the Devolution Initiative was that it would have taken too long and been very costly to bring in someone new to the initiative and to the Foundation to do the study. On the other hand, in answering what did the Foundation learn about the process of devolution in communities, and what was learned about how to co-evolve evaluation and grantmaking in the context of a cross-program initiative, Behrens believes that a meta-evaluation using a fresh look might have been more useful. "There was a small budget for this learning exercise, and one lesson is that we need to have enough to do this right, and spend the resources we do have on the right things," she said. (The LIL study for the Devolution Initiative was budgeted at \$50,000, which included both extensive review of the many documents associated with the Initiative and interviews with all key players.)

A second internal evaluation of the DI was undertaken in 2002 in response to questions from the Foundation Board about whether there was "evidence that there were differences between the states where the Foundation's Devolution Initiative was active and states where the Foundation did not conduct the Initiative."²⁰ This evaluation concluded that:

- Specific policies changed [in the DI states] to better serve people transitioning from welfare to work [compared to the non-DI states];

²⁰ This is the language of a "Program Note" in the materials for a December 2003 meeting of Foundation Trustees. The Program Note introduces the report resulting from the internal evaluation, which was entitled "Making a Difference in Devolution Initiative States: A Comparative Review."

- More people [in the DI states], especially those most directly affected, were significantly involved in devolution-related policy changes; and
- Grantee organizations themselves changed to better serve their constituents [in the DI states].

This study involved an analysis of data on key indicators of health coverage and the use of TANF funds for child care in the DI states and five comparison states as well as a survey of 23 state policymakers in both sets of states administered by the Center for Policy Alternatives, and a review of final DI evaluation reports.

Appendix A

Persons Interviewed for the Case

C. Patrick Babcock, former WKKF Director of Policy and DI Co-Chair

Teresa Behrens, WKKF Director of Evaluation

Julia Coffman, Harvard Family Research Project

Karen Lake, WKKF Director of Communications and Marketing

Mark Lelle, former WKKF Evaluation Manager and consultant to WKKF

Daniella Levine, Communication Health Action Information Network
(Human Services Coalition of Dade County, Florida)

Ricardo Millett, former WKKF Director of Evaluation

Dan E. Moore, WKKF Vice President for Program and DI Co-Chair

Anne Petersen, WKKF Senior Vice President

William Richardson, WKKF President

Barbara Sabol, WKKF Program Director, Health

Miguel Satut, WKKF Program Director, Youth and Education

Henrie Treadwell, former WKKF Program Director, Health

Alice Warner, WKKF Program Analyst

Heather Weiss, Director, Harvard Family Research Project



Evaluation Roundtable

Diane Smith, WKKF Program Assistant, provided invaluable assistance in gathering the documentation for this case, hunting facts, and correcting errors.

Evaluation Roundtable

Learning Meetings	Reports/Briefs
1997 December	Expectations, Implications, and Early Achievements
1998 June	Partnerships, Dissemination, and Capacity Building
1998 October	1998 Learning Report and Executive Summary
1998 November (no meeting)	Community Foundation Brief: Strengthening Prospects for Sustainability of State and Community Mobilization Efforts
1999 February	DI Evaluation Key Accomplishments and Future challenges Special Brief on State Rollouts
1999 June	Learning Meeting Memo
1999 November	Learning Meeting Brief Pathways Diagram
2000 March	Overall Devolution Key Accomplishments 1997-2000 State Mobilization Reports 1999-2000 (5 reports)
2000 June	Understanding Devolution Through Ethnic Media Citizen Research: Unpacking Pathway 3 Scholar Practitioner Program Brief
2000 August (Detroit meeting)	Memo on community mobilization framework
2000 October	The Devolution Initiative: State Views on Sustainability
2000 December	2000 Year-End Wrap-Up State Logic Models
2001 January	Five DI Lessons Linked to Sustainability DI Lessons Applied to the 5 States Memo on Tracking TANF Reauthorization
2001 March	DI Cross-State Media Strategies DI State-Specific Media Strategies (5 reports)
2001 June	Media Research Findings in Brief: 1999-2000
2001 August	2001 State Policy Advocacy Results: Cross-State Summary State 2001 State Policy Advocacy (5 reports)
2001 October (no meeting)	Scholar Practitioner Program Year 2 Evaluation (Revised)
2002 February	Learning Report 1997-2001 Evaluation Findings Pathway 1 (1997-2001) Findings and Implications Pathway 2 (1997-2001) Findings and Implications Pathway 3 (1997-2001) Findings and Implications
2002 June	“Post” Devolution Initiative Readings on Sustainability, TANF Reauthorization and State Policy Developing Community-Based Researchers (SPP)
2002 September	Scholar Practitioner Program Year 3 Evaluation Brief

2003 February	DI Sustainability: Cross-State Findings State Sustaining Devolution Initiative Work and Gains (5 reports) Ten Questions about Investments in Citizen Engagement
---------------	---

← Informed conversations on state mobilization

Helped to catalyze a sustainability conversation early on

←

← Introduced pathways framing

← Began reporting on five states separately

← Examples of special topic evaluation reports

← Clarified Pathway 2/3 distinctions

← Reports on sustainability specifically for State PDs

← Reporting on policy-related outcomes

← Final reports on full DI

} Post-Initiative follow up on sustainability and other topics

Appendix C
Devolution Initiative Evaluation Methods

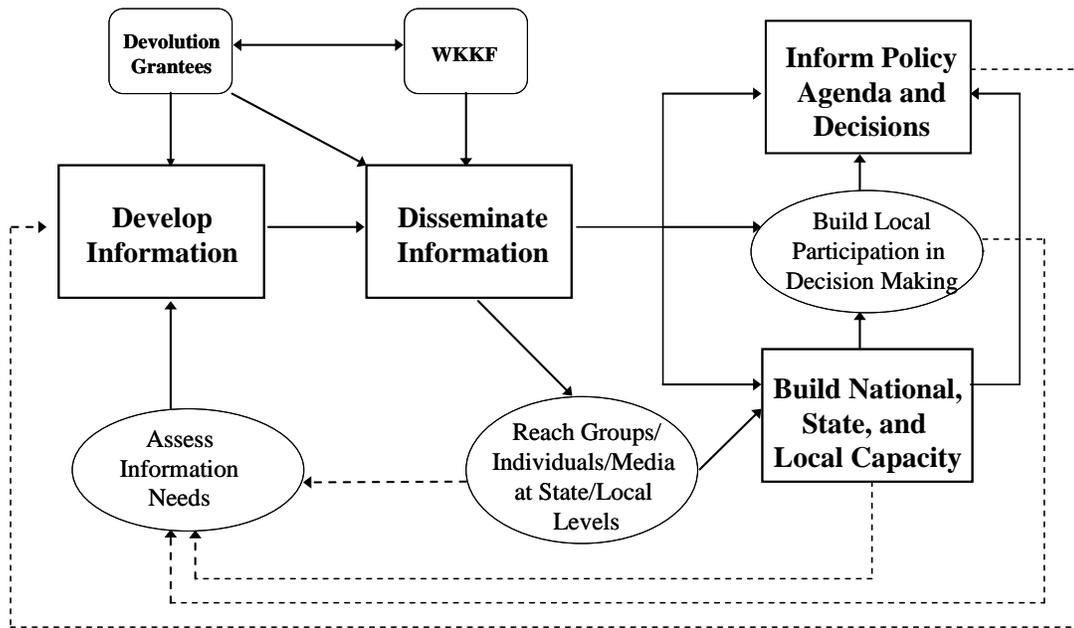
Method and Source	Data Amount & Time	Data Type & Analyses
National Grantee Document Review	129 in 1997 325 in 1998 311 in 1999 456 in 2000	Quantitative Content Analyses
National Grantee Web Site Review	19 Web sites in 2000 and 2001	Quantitative and Qualitative Content Analyses
National Grantee Questionnaires and Interviews	9 in 1998, 11 in 1999, 15 in 1999, 16 in 2000, 17 in 2001	Quantitative Statistics; Qualitative Content Analysis
State Grantee Questionnaire, Interviews, Written, and Electronic Communication	All 10 state grantees twice each year and across the five states in 1999, 2000, and 2001	Quantitative Statistics; Qualitative Content Analysis
Grantee Staff Interviews	15 field staff and organizers (2000)	Qualitative Content Analysis
Community-Based Partner Interviews and Questionnaires	32 community-based partners (2000)	Quantitative Statistics; Qualitative Content Analysis
Coalition Member Survey	203 coalition members (2001)	Quantitative Statistics
Scholar Practitioner Interviews	Two rounds of formal interviews with SP mentors from 1999-2001; participant observation at 3 annual conferences; informal communication	Qualitative Content Analysis
State Site Visits and Event Observation with Structured Protocol	~ 35 across five states in 1999-2001 (includes site visits, observation at events, cross-initiative meetings)	Qualitative Content Analysis
Policymaker Questionnaires and Interviews (legislators and administrators)	25 in 1998-1999 51 in 1999-2000 28 (interviews) in 2001	Quantitative Statistics; Qualitative Content Analysis
Policy Tracking	Policy alerts and electronic legislative tracking 1999 – 2001	Qualitative Content Analyses
New Voice Interviews	15 in 2001	Qualitative Content Analyses
Participant Observation in WKKF Meetings and Conference Calls	>150 from 1997 – 2001 (includes Core Team, networking, disseminator, consultant, state, cross-initiative)	Quantitative Descriptive Statistics; Qualitative Analysis
Descriptive Scan (memos, reports, email)	Series over 1997 – 2001	Qualitative Content Analysis
Media Tracking	2,391 articles about health care and welfare policy issues. From 29 electronic and print sources	Quantitative and Qualitative Content Analysis
Citizen Research	10 Rochester Meetings; 10 Middletown Meetings 1999-2001	Quantitative Statistics; Analysis of Audio Tapes; Debriefing

Evaluation Roundtable

Project-level Evaluation (including TA on meetings)	Reports from 3 project-level evaluations	Qualitative Review
---	--	--------------------

Appendix D

WKKF Devolution Initiative Logic Model



Appendix E

Case Teaching Questions

The purpose of a teaching case, such as the one about the Kellogg Foundation's Devolution Initiative presented here, is to raise issues for discussion that are typically met in situations of uncertainty and complexity, where uncertainty and complexity may be present in the tasks, the organizations, the methods and technologies, the wider environment, or all of these realms. To facilitate and stimulate discussion, teaching cases are constructed to provide their audience some basic facts about the situation, a narrative of events, and the perspectives of multiple participants in the events -- without conclusions or judgments.

Teaching cases are taught through the use of questions to uncover the views of participants in a case teaching session about the issues the case is intended to open up. The case teacher begins the session questioning participants about the basic facts of the case as presented in the case document in order to establish a common understanding – what is the situation, who are the players, what is the task, what is the past experience of the organization with similar tasks, what is the policy environment, what is the state-of-the-art, what is the evaluation design, etc.? Then, the case teacher turns to questions for which there is no correct answer. The following questions are offered to assist teachers of this case.

1. What was the larger context for “devolution”? What was going on at this time in politics, government, and society that led to Devolution?
2. How did the Foundation's mission lead to the Devolution Initiative? What rationale emerged to connect this Initiative to the Foundation's mission? In what ways was this a departure for the Foundation? How did this change occur? And how did this change affect the conceptualization of the evaluation?
3. How did the “Evaluation Premises” paper affect the evaluation framework?
Discuss these premises. In the larger field of evaluation, where do these premises position the W.K.Kellogg Foundation's approach to evaluation? What values are embedded in the premises?
4. How did the logic model change over time? What led to these changes? What was their importance?
5. How was the evaluation managed by the Foundation? What were the effects of this management approach? What are alternative ways of managing an evaluation?

Evaluation Roundtable

6. How were issues of attribution handled in the design? How did concern about attribution change over time? Why?
7. What were the evaluation's "multi-dimensional set of services to the Foundation"? What tensions are manifest between these various "components" of the overall evaluation?
8. Discuss the notion of an evaluation "bottom line" as it emerged in the evaluation? Where does this metaphor come from? What are alternative metaphors?
9. How did ethnicity and race become an issue in the evaluation? Discuss the issue of diversity in an evaluation context based on the case?
- 1-10. _____ What were the various roles played by the evaluators? How did these roles change over time? What tensions emerged between these various roles? What lessons do you take from these role examples?
- 2-11. _____ Do you concur or not in Pat Babcock's observation that, if he were going to do the Devolution Initiative again, he would bring in different evaluators for different stages of the project? Why?
- 3-12. _____ From the evaluators' perspective, what are the advantages and disadvantages of having an internal audience or an external audience or both? From the Foundation's perspective, what are the advantages and disadvantages?
- 4-13. _____ Why was there no overall, "summing up" final report from the evaluators? What were the results of this inside and outside the Foundation?
- 5-14. _____ What does it take – from both evaluators and stakeholders – to ensure that learning takes place? What ingredients and commitments are necessary to ensure an atmosphere conducive to learning? What forms and formats did evaluation reporting take in the DI case and how did these contribute (or not) to learning?
- 6-15. _____ How can interest and engagement in evaluation be sustained throughout an initiative – especially over time as players change and the Foundation adopts new priorities? Are there ways in which the evaluators and the Foundation could have helped to sustain interest and engagement longer in the Devolution Initiative?