Program evaluation is a form of applied research—a systematic, data-based process for judging the value of a program, helping to make decisions, or creating information about key activities or processes. During the past 40 years, the field of program evaluation has developed simultaneously in two directions. One branch has focused on accountability, requiring proof that money has been well spent and that staff have implemented programs with fidelity. This approach typically requires outcome measurement, performance assessment, or cost data, providing summative evaluation information for audiences beyond agency staff or administrators. Another branch of evaluation has focused on collecting information that will enable staff (and sometimes clients themselves) to improve programs. This approach focuses on collaborative inquiry and may couple professional evaluators with program staff or participants to engage in ongoing, bottom-up, participatory activities that, through the acts themselves, teach people evaluation skills. In its fullest form, the evaluator’s role changes from that of outside expert to coach and quality control manager, guiding organization members in their own evaluative work.

Because external funding in the form of grants and contracts routinely comes with accountability strings attached, many administrators have accepted the self-evaluation challenge, and staff in agencies and schools regularly engage in annual improvement processes involving data collection and analysis. However, the focus on accountability has also created serious problems for such organizations. First, the demand for evaluation information has created many “accidental evaluators”—individuals with little or no training in program evaluation who are nonetheless required to conduct evaluation studies. In general, few practitioners have formal training in evaluation procedures; many exhibit negative attitudes toward program evaluation; and most would prefer to serve...
additional clients rather than devote precious fiscal resources to evaluation. The demand for program evaluation also has created a dilemma for not-for-profit agencies: They must provide accountability evidence that satisfies funders and policy makers, while at the same time generating information that can help staff do a better job serving people. These two types of evaluation information are rarely the same and, when push comes to shove, outcome requirements typically dominate evaluation activities. Where there are limited resources available for the work, outcome measurement may unavoidably overwhelm formative efforts.

Evaluation capacity building (ECB) is a fairly recent conceptual development that attempts to address some of the problems inherent in program evaluation. Although the practice is sometimes considered distinct from program evaluation, the goal of ECB is to strengthen and sustain effective program evaluation practices by increasing an organization’s capacity to

- design, implement, and manage effective evaluation projects;
- access, build, and use evaluative knowledge and skills;
- cultivate a spirit of continuous organizational learning, improvement, and accountability; and
- create awareness and support for program evaluation and self-evaluation as a performance improvement strategy in the internal and external environments in which they function.

Building on the collaborative inquiry approach, ECB is halfway on a continuum, with user-focused evaluation at one end (where an evaluator acts to increase the likelihood that someone will use the results of a study) and organization development or continuous quality improvement at the other (where people within an organization engage routinely in data collection, reflection, planning, and action). Evaluation capacity building simultaneously addresses demands for accountability and for interactive participation. In theory, it is both timely and cost-effective.

Our study, which was conducted during 2003 and 2004 with support from a CURA Faculty Interactive Research Program grant, had a dual purpose: first, to examine the development and overall status of program evaluation in three Twin Cities not-for-profit organizations that had a long-term interest in the process; and second, to study the viability of evaluation capacity building as a policy outcome that could generate quality data in a timely and cost-effective manner. This article will briefly outline the study’s methods, identify the common features of program evaluation across the three organizations we studied, present a grounded framework for evaluation capacity building, and consider the implications for institutionalizing program evaluation in organizations large and small.

Study Methodology

We used three criteria to identify potential organizations to study: (1) The organization had to be a Twin Cities-area not-for-profit organization or school district with an external mandate for accountability information; (2) there had to be an ongoing, routine program evaluation function in the organization; and (3) one or more high-level administrators in the organization had to be committed to institutionalizing or broadening the evaluation function. Many agencies and districts easily met the first two criteria, but the third criterion narrowed our subject pool. We ultimately selected three organizations where one or more leaders understood the potential of program evaluation and were either interested in institutionalizing the process more broadly or were actively doing so: Neighborhood House, a social service agency that began as a settlement house for immigrants more than 100 years ago; the Science Museum of Minnesota, a community resource to teach science to Minnesota’s citizens informally; and Anoka-Hennepin Independent School District (ISD) 11, which is rapidly becoming one of the largest school districts in the state.

Although they differ on many dimensions, the three organizations we chose to examine have certain similarities. They are all large organizations with a governing board and a bureaucratic hierarchy; they are well-known community entities, having existed for at least 50 years (and in two cases longer); they have an educational mission that includes a commitment to serving society; and they share the collective challenge of conducting program evaluations and demonstrating to their boards, funders, and staff the value of their work, their efforts to improve activities, and their achievement of outcomes.

At the same time, because we wanted to study program evaluation in distinct contexts, we chose organizations that differed from each other in terms of mission, activities, clientele, and structure. Not surprisingly, their evaluation structures differed as well. At the time this research was conducted, the Science Museum had no one designated as the internal program evaluator, and staff either hired external consultants or conducted evaluations themselves.1 Neighborhood House had a full-time internal evaluator who coordinated certain activities, but staff hired external evaluators for some projects or were responsible for completing others themselves. Owing to its size, tradition, and the accountability mandates it faced, Anoka-Hennepin ISD 11 had no single individual with the role of program evaluator, but rather had a number of people connected to its evaluation function: a student assessment office with a full-time professional and three staff members, other staff who worked on mandatory and targeted evaluation projects, and external evaluators who completed evaluation contracts. Our intention was to study the status of program evaluation and the prospects of evaluation capacity building across these organizations, seeking commonalities and identifying the components of successful evaluation capacity-building efforts.

During the study, we were participant observers in more than 25 evaluation-related meetings across the three organizations, sessions in which people framed evaluation questions, worked on instruments, interpreted data, reacted to reports, and so on. We collected organizational documents for analysis, including evaluation reports, instruments, descriptive materials, and accountability mandates. Finally, we conducted a series of formal interviews with more than a dozen leaders or evaluation champions, supplemented by informal interviews with numerous other individuals engaged in evaluation activities.

Commonalities of Program Evaluation

Despite the differences in evaluation structure across the three organizations, the status of program evaluation within them was fairly similar. Absent leadership and a purposeful focus on doing something different, mandated evaluations clearly drove the evaluation agenda in all three organizations. Not surprisingly, respondents reported that

1 The museum hired an internal evaluator in the spring of 2005.
external grants, an important source of funding for each organization, typically required evaluations. As an administrator from Neighborhood House put it, “The circumstances under which [program evaluation] is done is often in response to funding requests or reports to funders. . . . They are mandated.”

A program director from the Science Museum noted, “It is accountability to the funder, but not to the museum.”

Given federal funding requirements, an Anoka-Hennepin ISD 11 coordinator said, “We need to really understand where our kids are at and the impact of [state tests] on our students and the funding for [federal programs]. It’s become much more on the radar screens since No Child Left Behind [a federal testing requirement] came through.”

If mandated evaluations were the first commonality, a second commonality made it difficult for staff to respond to such mandates. In each of these organizations, the key internal resource for program evaluation—time—was extremely limited, with people reporting few available hours for evaluation activities. Comments from across the organizations document this frustrating similarity. In the words of one museum administrator, “Whatever it is that we come up with investing more time on evaluation, I honestly don’t know . . . [H]ow are we going to juggle things around so that we can give time to evaluation?” Another museum administrator added, “We don’t have a lot of evaluation resources, so this comes on top of everything else that program directors, program managers do, and their staff.” An Anoka-Hennepin ISD 11 official observed that “evaluation resources basically come and go . . . definitely, with the wind,” and a colleague concurred: “The biggest challenge is resources, which is always money, because money will buy time, money will buy people; funding, budget will buy materials, access, training, everything you need to change and to implement [program evaluation] and to have it become systemic.” A top administrator at Neighborhood House lamented, “Because of our growth over the last couple of years, it is so easy to get caught up in your own programs. . . . [W]e have so much work that you don’t ever step back and breathe and see the big picture.”

With the reported pressures of mandated evaluations and the lack of time to conduct them, the viability of evaluation capacity building in organizations like these seemed an obvious concern. Happily, the efforts of staff and administrators in the organizations we studied pointed to ways that capacity can be built, slowly and systematically over time, through procedures that make sense even for small organizations. The next section describes the framework that emerged from our case study data, an outline that organizations can use in their efforts to build evaluation capacity.

A Grounded Framework for Evaluation Capacity Building

What we learned from the three organizations we studied helped us develop a conceptual framework for understanding and developing ECB. Based on the experiences of evaluation champions in these organizations—and each organization had several—the framework provides a common and consistent approach to developing a practical evaluation function when planning and implementing organizational performance improvement strategies and accountability mechanisms. It is intended as a resource for a wide range of stakeholders in not-for-profit organizations seeking to increase their long-term capacity to conduct and use program evaluations in everyday activities. This framework encompasses findings from work in different settings and, with minor adjustments, can cater to the needs and circumstances of both large and small organizations. Although the framework is focused on not-for-profit organizations, we believe it also provides a viable approach for other organizations interested in improving the quantity and quality of their evaluations.

The evaluation capacity building framework we outline here consists of three major categories—organizational context, ECB structures, and resources—each of which we will describe using examples. Some components of our framework may at first seem rather generic, and there is a simple reason for this. Evaluation capacity building has been developed through research from several intersecting fields, notably evaluation studies, human resource development, organization development, adult learning, and social and industrial psychology. Accordingly, the components of the framework are derived from the theories and practices of these participating disciplines. One of the strong points of the ECB framework is that, while unequivocally centered on promoting and institutionalizing evaluation in organizations, it is also applicable to a number of other organizational interventions—for example, creating a
Organizational Context. As shown in Figure 1, organizational context consists of two components. **External** organizational context locates an organization in time and place, and itself consists of two components: (1) external mandates and other requirements stemming from required accountability measures (e.g., program or project evaluations associated with grant funding or federal or state reporting requirements), and (2) an external environment supportive of change (e.g., a professional community interested in evaluation processes or accreditation agencies that encourage innovation). The mandate to evaluate (for accreditation, legislated accountability, or grant requirements) has been key to continuing evaluation efforts successfully. Nowadays, ECB practitioners must capitalize on societal requirements for accountability without letting accountability completely define the process of building evaluation capacity in organizations. For instance, the National Science Foundation funds a number of projects at the Science Museum of Minnesota with the specific stipulation of a formal evaluation plan from the program’s inception. Similarly, public and private sources concerned about the impact of their investment fund several Neighborhood House programs.

The **internal** organizational context is key to determining the feasibility of ECB. The three organizations we studied suggest that a positive, ECB-friendly internal organizational context has five components, each important to ECB efforts. Creating this internal context, however, is not an evaluation “silver bullet” that will succeed in all settings.

The first component we identified is supportive leadership that shares responsibility for ECB. Leadership is key to capacity building. In its absence, the routine demands of mandated studies and ongoing activities eliminate the possibility of capacity building. Apart from fiscal support, administrators and opinion leaders can support the ECB process by providing verbal support in public (for example, when the superintendent at Anoka-Hennepin ISD 11 hosted an evaluation meeting) and by serving as role models who evaluate and monitor their own activities. Leadership at Neighborhood House, for example, had made a commitment to building evaluation capacity several years earlier. As one administrator put it, “There’s a much better understanding of the need for evaluation. There’s upper management support for getting evaluation done and what that may entail.” Led by a superintendent with a background in research, the Anoka-Hennepin ISD 11 leadership similarly strongly endorsed ECB efforts.

The second component is an evaluation champion. The value of evaluation champions is hard to overestimate. Although leadership is critical, so too are the activities of people who champion the evaluation cause. The highly visible and engaging “Evaluation Platoon” at Neighborhood House is a good example of people voluntarily implementing evaluation activities and purposely passing their evaluative spirit and knowledge to their colleagues. Science Museum staff recognized the significance of evaluation supporters in mainstreaming evaluation. As one administrator pointed out, “Some of our program staff . . . actually are conducting evaluations and getting excited about it and telling others, so it’s that sort of peer-to-peer building of the infrastructure that says this is a place where evaluation is an important part.”

The third component to building effective evaluation capability is broad-based interest in and demand for evaluation information. Until organizational leaders have a vision of what evaluation can do and support is built among the rank and file, evaluation capacity building will not be fruitful. Evaluation capacity building requires a broad stakeholder base and a special concern for the representation of the ultimate intended beneficiaries. At the Science Museum, members of the management team and staff thought seriously about how to build their capacity to mine more information from existing evaluations and about how to ensure that what they learned was shared more broadly and consistently across the institution. One manager suggested an evaluation committee that would meet regularly to look at evaluation results and think about the kinds of cross-cutting questions important for a museum to investigate. Even young beneficiaries of certain museum programs got involved meaningfully by creating their own evaluation questions about their projects and collecting data from visitors and from each other.

A fourth component, linked to the internal context of an organization,
is an internal environment supportive of change. Such an environment is characterized by the following favorable conditions for successful evaluation capacity development: open mindedness, lack of fear (of being penalized), respect for each other, rewards for innovation/risk-taking/creativity, a sense of humor, and positive attitudes toward evaluation. In the ECB process, there is a strong need to develop a receptive culture in which demand for and effective use of evaluation output can grow.

In Anoka-Hennepin ISD 11, district leadership has worked for several years to involve a wide range of staff and stakeholders in evaluation activities, knowing that their meaningful participation in answering evaluation questions thoroughly is more important than getting results that look or feel good.

The final component of the organizational context category, sufficient input in decision making, simply means that people in the organization must be able to use data to make decisions. This opportunity, according to one person we interviewed, helps his organization “to do a better job of assembling and corralling all of these ideas such that we decide which project ideas have the highest merit in terms of delivering on our audience goals, our learning goals, [and] our financial goals.”

**ECB Structures**. Like the first category, the second category of the evaluation capacity building framework, ECB structures, also consists of several components. This category has direct implications for those interested in building evaluation capacity because interested people can actively develop these structures within organizations, purposefully creating mechanisms to build evaluation capacity. These structures (see Figure 1) were either present in the three organizations we studied or, if not present, were acknowledged within the organization as necessary for moving the ECB effort forward.

The first component of ECB structures is a *purposeful ECB plan for the organization*. This implies an appropriate conception of and a tailored strategy for evaluation in organizational policies and procedures (aligned with the organization’s mission, goals, and strategies); an evaluation oversight group (for example, the evaluation manager and Evaluation Platoon at Neighborhood House); and a formal ECB written document.

A second component of ECB structures is the *infrastructure to support specific components of the evaluation process*. At minimum, this includes a question-framing mechanism to generate evaluation studies; a system to measure organizational needs; the capacity to create evaluation designs and collect, analyze, and interpret data; the presence of an internal reporting, monitoring, and tracking system; and public relations capability. The evaluation infrastructure might include such mechanisms and techniques as satisfaction surveys, focus group procedures, case analyses (strengths and weaknesses), statistical analyses, benchmarking, quality circles or problem-solving groups, chart reviews, and periodic monitoring.

Clearly, an internal evaluation and learning system that an organization has already developed to track and reflect on its performance should be reinforced and built on. In one Science Museum youth program, for example, an informal internal system was converted and formalized to track activities and reflect on project performance more systematically. There were four components of the eventual system: journaling, an attendance tracking system, digital portfolios, and demonstrations.

Another component of ECB structures is *purposeful socialization into the organization’s evaluation process*. The core ingredients of this socialization process are clear expectations of evaluation roles (that is, the notion that everyone is expected to “do” evaluation); incentives for participation; formal training, professional development, or coaching in evaluation; and learning evaluation by doing it. In the ECB process, the focus is on working with the whole organization, not just on developing the skills of individuals, by building awareness of techniques and approaches that are workable in associated contexts and by developing evaluation skills appropriate for each level of an organization through a range of training opportunities.

Literature and our own experiences have highlighted the importance of linking more formal training with “hands-on” experiential learning as an approach to building evaluation capacity. The development of evaluation capability in the organizations we studied has been promoted through customized evaluation trainings (for example, a specialized course for Neighborhood House staff, and voluntary evaluation seminars for staff of the Science Museum of Minnesota), as well as “learning by doing” activities (for example, stakeholders’ involvement in defining questions, data collection, and methods). For instance, administrators and participating teachers in Anoka-Hennepin ISD 11 took part in several evaluations by defining evaluation objectives, questions, methods, uses, and so on. Similarly, Neighborhood House staff learned valuable skills in...
refining program goals and objectives, developing logic models, establishing standards and indicators, and developing and administering surveys.

Evaluation capacity building work should include meaningful incentives for participation in evaluation—for example, stipends and honoraria, staff recognition within the organization, and (especially) the time and flexibility necessary for people to learn to integrate the evaluation process into their ongoing work. In an interview, one Neighborhood House administrator told us about trying “to say to folks that this is important and critical work” and that this “is why I joined the evaluation Task Force and team. . . . It also sends a message to the staff that if I can make it happen in my calendar, they need to do also.” Research shows that the most potent motivators for adult learning are internal. In this case, it is important to emphasize to learners that participation in the evaluation process can lead to the development of valuable, lifelong evaluation/research skills and the capacity for self-critique, self-determination, and systematic inquiry.

Peer learning structures are the final component of ECB structures. Purposeful trust building and interdependent roles in an organization should ensure that collaboration is present over time (not just for a one-shot effort) and is not initiative-driven. Distrust of the evaluation process is surprisingly common, even in an organization with good intentions. In Anoka-Hennepin ISD 11, building-level staff in one evaluation study at first asked how the evaluation would affect them, who would see the data, and whether these evaluation activities would be “one more thing that is dropped after a year or two.” Real trust, both interpersonal and organizational, can only be established over time, and it remains a fragile commodity, especially in a large organization.

There is a need to be sensitive to the perspectives held by varying non-evaluator participants and to develop strategies to boost interest and engagement. There should be ample opportunities for reflection, for example, as well as open discussions of successes, challenges, and failures. The ECB effort also must allow adequate time and opportunities to collaborate, including, when possible, being physically together in an environment free from interruptions. A feedback mechanism in the decision-making process and an effective communication system are crucial in increasing the likelihood of cooperation in the ECB process. For instance, Anoka-Hennepin ISD 11 has institutionalized a practice of summarizing key messages from important meetings and creating standardized “action minutes” that document meeting content, decisions, and who will do what next.

Resources. The final category of the ECB framework acknowledges the reality that without resources, the work of capacity building must be taken from someone’s hide—a difficult sell when staff may already feel overworked. There are two aspects to the issue of resource needs. First, organizations must have easy access to evaluation resources. These resources include (1) formal training or professional development in evaluation (which may also be included in the socialization process); (2) formal and informal just-in-time evaluation coaching; (3) personnel in the form of internal professionals and external consultants; (4) relevant research bases that contain “best practice” content; and (5) information on resources for evaluation, including books, journals, and online resources.

The second area of resource need is explicit sources of support for program evaluation in the organization. This support can take several forms. Fiscal support from the board or administration is both a signal and a tool. It includes basic resources (for example, copying, computer hardware and software, and the means to conduct data analysis), as well as explicit, dedicated funding for program evaluation activities. As a Science Museum vice president said, “We’ve got all of these projects; each one carries thirty to fifty thousand more dollars over the course of that project for evaluation. You add those pieces together, you’ve got some real dollars.” A Neighborhood House administrator noted, “We’ve committed the financial resources to support [evaluation]. . . . Five percent of program revenue goes to support program evaluation and effectiveness.” In that agency, every grant or contract written builds resources for evaluation and supports evaluation activities. Resources can also include time within the workday to collaborate on evaluation activities, a luxury for practitioners like social workers or classroom teachers. Revenue-generating strategies to support ECB (for example, selling materials or having staff provide evaluation consulting in other agencies) may be a creative way to enable the evaluation process to sustain itself.

Conclusion
Evaluation capacity building is an example of a system of guided processes and practices that necessarily includes a wide variety of adult learning processes. Psychologist Carl Rogers once noted that learners want to be problem solvers, and adults are no exception; they tend to demand an immediate application of their learning. Evaluation capacity building builds on this need by requiring the active, self-directed participation of learners during the entire evaluation process. Applying the ECB framework, people learn evaluation by doing it themselves in their own programs and organizations. It is a context-dependent, learner-dependent, and learner-centered intentional action system. The wide-ranging backgrounds of participants must be taken into account, and learning materials and activities should allow for different types of previous experience. A number of lessons from the principles of adult learning are applicable to ECB, including establishing a relaxed, trusting climate conducive to learning; involving learners in an assessment of learning needs; mutual planning; and flexibility.

Why evaluation capacity building, and why now? One administrator at Neighborhood House nicely captured the importance of evaluation in his organization:

If we cannot say clearly and effectively who we are, what we do, and how effective we are, we’re not going to be able to continue to support programming or to take on new programming initiatives. It’s a way to say to people, here’s why you can believe what we are saying. It’s not just the numbers, it’s what happened in people’s lives. It’s the stories that go on with people.

By developing ways to help organizations assess their work, examine what is working and what isn’t, and learn how to strengthen program activities and increase their impact, building the evaluation capacity of organizations can be of paramount significance. At the end of one interview, a study participant who is also an avid gourmet cook compared program evaluation to a measuring cup, noting that evaluation helps us measure what we do, and you fill it and empty it . . . . It reminds me [to] test it, try it, taste it, does this work, this doesn’t work. How do you adjust? But something
great comes out of it if you work at it. Something great comes out of it.

Our study examined three organizations facing the central program evaluation challenge of this first decade of the new millennium: responding to accountability demands, while at the same time generating information useful for program improvement. The accountability demands were easy to document, as was the amount of time needed for organizations to accomplish this work—time that is rarely available in not-for-profit agencies or school districts. Our second purpose was to study the viability of evaluation capacity building as an outcome that addresses these dual concerns by generating quality data on an ongoing basis as part of the day-to-day functioning of the organization.

The grounded framework that emerged from our data outlines what is needed to develop ECB and is available as a template for any organization interested in mainstreaming the evaluation process. Even organizations with limited resources for this work can focus on creating structures that will, over time, support evaluation activities incrementally. Organizational leaders can develop an ECB plan and establish peer learning opportunities on a small budget and then systematically work to socialize long-time staff and newcomers alike to evaluation as a way of life. Minimal incentives—especially in the context of externally mandated evaluations—may encourage people to collaborate on evaluation activities that, taken together, can begin to foster evaluative thinking across programs. Resources can surely help with this process, but they do not necessarily guarantee its success. The two-fold challenge is first to begin and then to sustain the evaluation process using available resources. The most profound lesson we learned about building evaluation capacity from the three organizations we studied, however, is also the most straightforward: Under the right circumstances, people and organizations can learn evaluation by doing it. That, we believe, is a lesson worth knowing.

Jean A. King is a professor in the Department of Educational Policy and Administration at the University of Minnesota, where she serves as Director of Graduate Studies and Coordinator of the Evaluation Studies Program. Her research interests include participatory evaluation and evaluation capacity building, program evaluator competencies, and interpersonal skill development for evaluators. In 2005 she received a College of Education and Human Development Community Service Award for her volunteer work at Neighborhood House. Boris Volkov is a Ph.D. student in the Evaluation Studies program in the Department of Educational Policy and Administration at the University of Minnesota. His interests include evaluation capacity building in organizations, the evaluation of educational technology and science projects, and human resource development in organizations.

This study was supported by a grant from CURA’s Faculty Interactive Research Program. The program was created to encourage University faculty to carry out research projects that involve significant issues of public policy for the state and that include interaction with community groups, agencies, or organizations in Minnesota. These grants are available to regular faculty members at the University of Minnesota, and are awarded annually on a competitive basis.

21st Annual Conference on Policy Analysis

The 21st Annual Conference on Policy Analysis will be held Wednesday, October 19, 2005, at the Continuing Education and Conference Center on the University of Minnesota’s St. Paul campus. The theme of this year’s conference is “Meeting the Challenges of a Changing Minnesota.” The conference is intended to provide an opportunity for analysts and policy makers to gain insight into current trends and changes in the policy-making environment, explore emerging policy issues, and share ideas with policy analysts from around Minnesota.

This year’s keynote speaker is Ember Reichgott Junge, attorney with The General Counsel, Ltd., and radio host for AM 950/VoiceAmerica Channel. The conference will also feature a panel discussion on tax policy and its effects on Minnesota’s economy, as well as concurrent sessions on policy research in nonprofit organizations, the role of local governments, Minnesota’s academic achievement gap, the I-394 MnPASS Corridor, the changing role of news media, the St. Croix River Basin restoration, and long-term care financing.

The conference is sponsored by the Economic Resource Group (ERG), a State of Minnesota consortium to promote the sharing of policy information. The registration fee is $125 if postmarked by October 5, or $140 if postmarked after October 5. Full-time students may register for $25 with a current class enrollment statement. For further information, visit www.cce.umn.edu/policyanalysis, e-mail conferences5@cce.umn.edu, or call Katie Kjeseth at 612-624-3708.